

# The Virtual Solar Observatory - A Component of the NASA Heliophysics Data System

Frank Hill and the VSO Team

# The Core Team

- Rick Bogart, *Stanford*
- Alisdair Davey, *CfA*
- Joe Gurman, *NASA GSFC*
- Keith Hughitt, *NASA GSFC*
- Joe Hourclé, *NASA GSFC*
- Piet Martens, *MSU*
- Kevin Reardon, *Arcetri*
- Jennifer Spencer, *Stanford*
- Igor Suàrez-Sola, *NSO*
- *Plus many others in the past and now working on services outside the “small box”*

# Outline

- History
- Design & Interface
- Current status
- Interoperability
- Lessons learned

# History

- Early “local” data centers
  - NASA/GSFC SDAC – 1980s
  - NSO Digital Library – 1993
- Original idea for synthesis: The Whole Sun Catalog (1995-1997) – Fleck, Bentley & Sanchez
- Funding attempts 1999-2001, successful in 2001
- First release late 2004
- Funding level: ~\$300k/year, total so far of ~\$3M

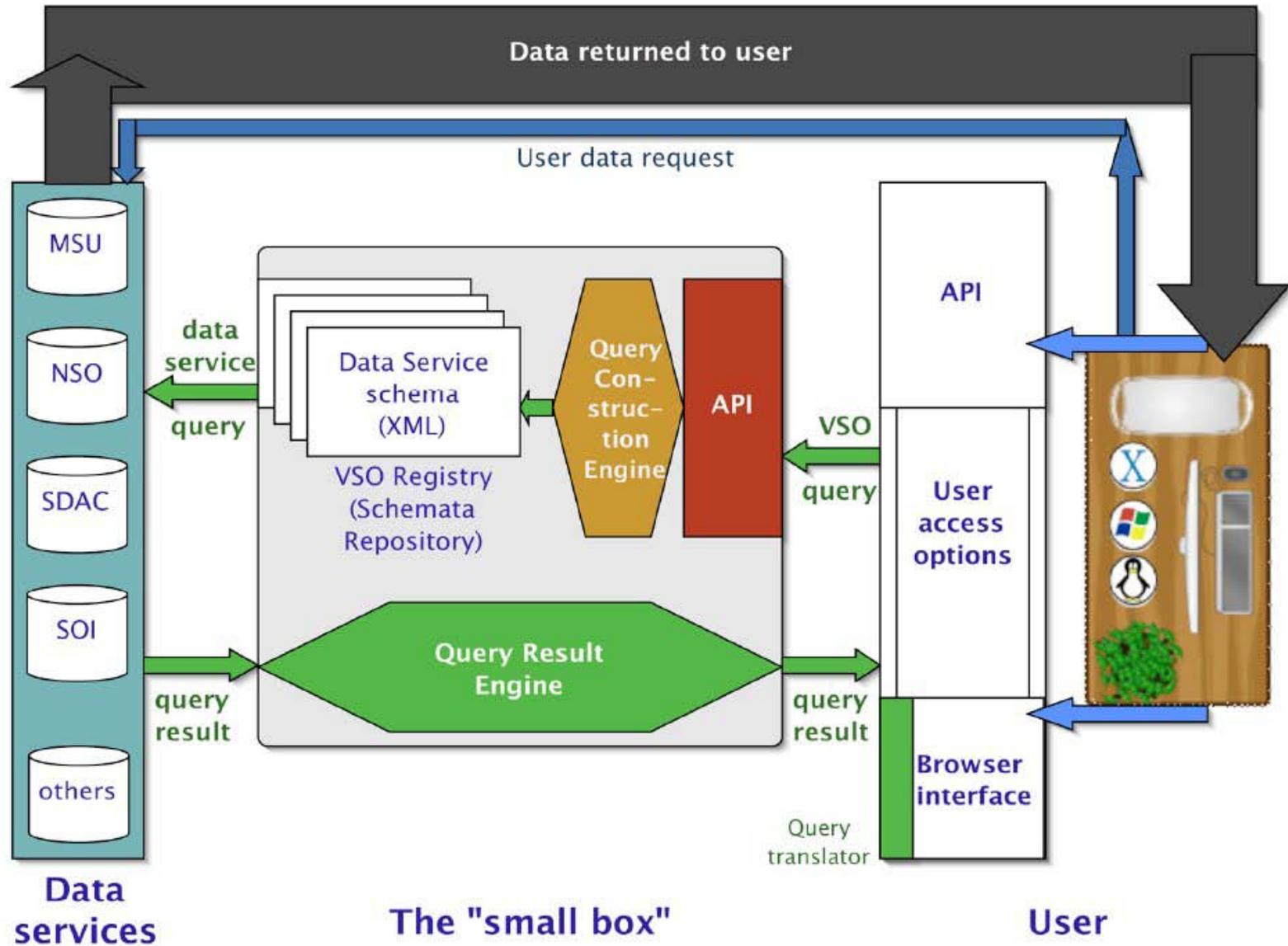
# Design requirements

- Distributed
- Search for and access multiple missions, observatories, and wavelengths without intimate knowledge of the data organization (*e.g. by physical observable and/or mission/ instrument*)
- Provide access to analysis software, instrument descriptions, &c. that enable use in scientific research
- Given the funding level, had to draw a “small box” around a small set of attributes useful for doing science

# What we decided we couldn't offer, at least at first

- Data mining (knowledge discovery)
- Remote processing: left for EGSO, CoSEC
- Grid computing (as opposed to data grids)
- Access control (ever)
  - We wanted only to be in the business of helping data providers to give data away, not preventing its access

# Resulting Design



# Multiple VSOs

- Because the VSO is a lightweight “small box”, there are several instances of the system
- “Official” instances are located at NASA/GSFC, NSO and Stanford
- Others are at CfA, MSU, ROB
- An individual user can have his/her own VSO on a local system
- URLs:
  - <http://vso.nso.edu/>
  - <http://vso.stanford.edu/>
  - <http://sdac.virtualsolar.org/cgi/search>



# VSO Time / Nickname Search Form

Version 1.4



Start Date/Time: 2007 Sep 01 / 21 : 00  
 End Date/Time: 2007 Sep 02 / 00 : 59  
 All Month All Day

Search Clear

## Nickname

**Note:** Nicknames generate an intersection with other search terms, so searching for a nickname, and a physical observable (or other parameter) when a nickname defines other physical observables will result in no matches.

### Dopplergram

- Full-disk dopplergram
- K-7699 dopplergram
- Na-D dopplergram
- Ni-6768 dopplergram

### Image

- 10.7cm image
- Ca-K image
- Coronagraph image
- EUV image
- H-alpha image
- Hard X-ray image
- He 10830 image
- Na-D image
- Soft X-ray image
- UV image
- White-light image

### Magnetogram

- Full-disk magnetogram
- LOS magnetogram
- Vector magnetogram

### Spectrum

- Atlas Spectrum
- EUV Spectrum
- IR Spectrum
- UV Spectrum
- Visible Spectrum

### Other

- Helioseismic Time Series
- Light Curve Time Series

Search Clear



Search VSO Help or enter Cart Id:

[VSO Glossary](#)  
[VSO FAQ](#)  
 Click on the icons for online help.

#### Query Menu [\[hide\]](#)

- [New Search](#)
- [Edit Search](#)
- [Continue Adding to Cart](#)
- [Click & Bookmark](#)
- [Email This Cart](#)
- [Track Your Request](#)
- [Back to Cart](#)

Search Status [\[show\]](#)

# VSO Time / Spectrum Search Form

Version 1.2



Start Date/Time: 2007 Sep 01 / 05 : 00  
 End Date/Time: 2007 Sep 01 / 08 : 59  
 All Month All Day

Search Clear

## Spectral Range

- soft X-rays [1 - 100 Å]
- extreme UV [100 - 1000 Å]
- ultraviolet [900 - 3800 Å]
- visible [3500 - 10000 Å]
- radio [0.3 - 30 GHz]
- OR select spectral range:

min   
 max   
 unit

Search Clear

## Notes

- Observable classification is tentative, as some data services have not registered full information on the classes of observables available.
- Time ranges of instrumentation provide the minimum and maximum ranges of data known to be available. Lack of an end date means that the archive is still receiving new information, but some archives may be a week or more behind the present date.

VSO @ [Home](#) | [NSO](#) | [Stanford](#)

Automatically Generated at : Fri Sep 28 21:18:15 2007

# CART ID: VSO-NSO-071012-046

## CART Data Request

Sessions : 12-Oct-2007 20:18:42 UTC

### Provider

SDAC

### Select Transfer Method

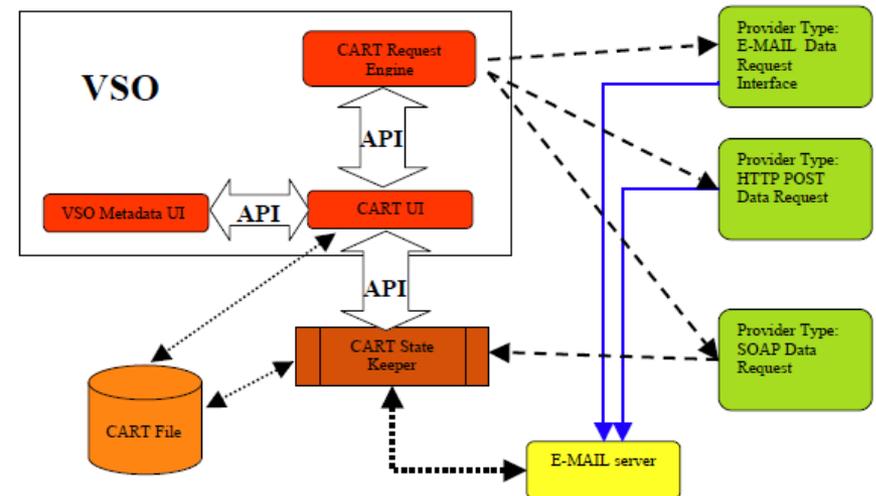
URL-FILE

STAGING-TAR\_GZ

# The Data Cart – A key concept

- Allows user to store and later exactly reproduce a search
- Carts are archived
- Can be cited in papers to allow others to replicate results

## Shopping VSO Style!



# The only VSO-supplied data service



**CART ID: VSO-NSO-071012-046**  
**VSO Movie Player**

Search VSO Help or enter  
Cart Id:

[VSO Glossary](#)

[VSO FAQ](#)

Click on the  icons  
for online help.

Query Menu [\[hide\]](#)

[New Search](#) 

[Edit Search](#) 

[Continue Adding to](#)

[Cart](#) 

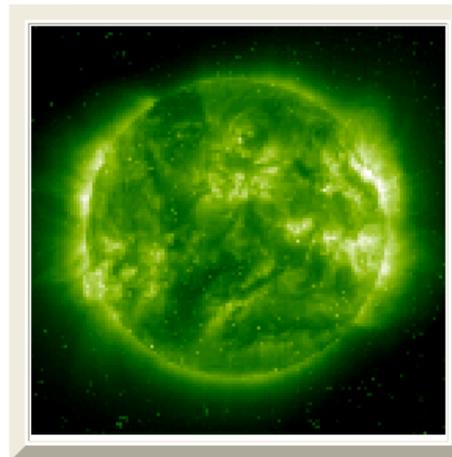
[Click & Bookmark](#)

[Email This Cart](#)

[Track Your Request](#)

Search Status [\[hide\]](#)

No Errors; No Warnings



Frame Start:  Stop:  Reverse:  Swing:

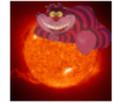
[Comments? Help us improve VSO](#) 

# Status

- Operational for more than 5 years
- More than 60 solar data sets going back to 1915
- Usage statistics
- Next steps



# VSO Time / Instrument Search Form



Version 1.2

Start Date/Time: 2010 Apr 20 / 06 : 00

End Date/Time: 2010 Apr 20 / 09 : 59

All Month All Day

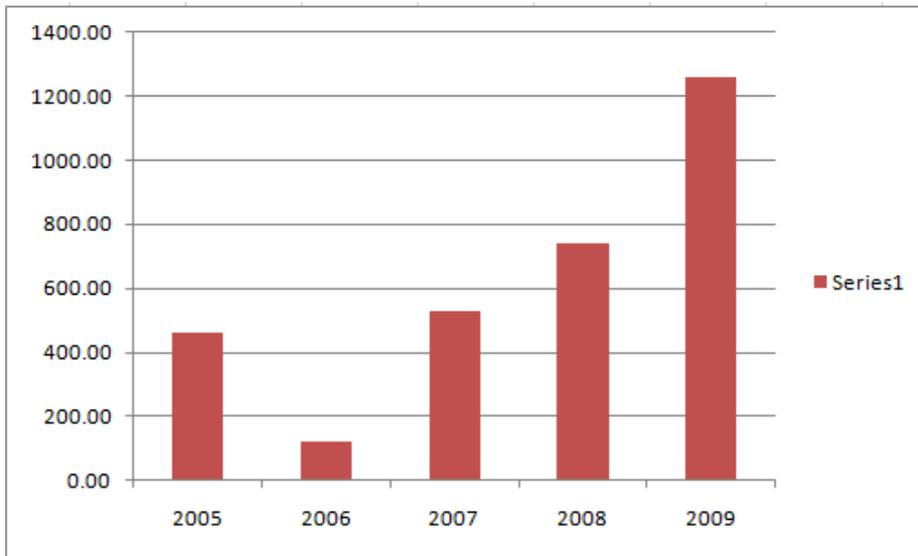
Search Clear

All from Provider	All from	Source	Instrument	Date Range
<input type="checkbox"/> HANET ⓘ	<input type="checkbox"/>	BBSO ⓘ	<input type="checkbox"/> BBSO ⓘ	2000.07.05 →
	<input type="checkbox"/>	KANZ ⓘ	<input type="checkbox"/> KANZ ⓘ	2001.02.07 →
	<input type="checkbox"/>	OACT ⓘ	<input type="checkbox"/> OACT ⓘ	2002.02.26 →
	<input type="checkbox"/>	OBSPM ⓘ	<input type="checkbox"/> OBSPM ⓘ	2004.10.22 →
	<input type="checkbox"/>	YNAO ⓘ	<input type="checkbox"/> YNAO ⓘ	2000.11.27 →
<input type="checkbox"/> HAO ⓘ		MLSO ⓘ	<input type="checkbox"/> chp ⓘ	1996.04.20 →
			<input type="checkbox"/> dpm ⓘ	1994.02.20 →
			<input type="checkbox"/> mk4 ⓘ	1998.10.01 →
<input type="checkbox"/> LSSP ⓘ		RHESSI ⓘ	<input type="checkbox"/> RHESSI ⓘ	2002.02.12 →
<input type="checkbox"/> MSU ⓘ		YOHKOH ⓘ	<input type="checkbox"/> BCS ⓘ	1991.09.01 – 2001.12.14
			<input type="checkbox"/> HXT ⓘ	1991.09.03 – 2001.12.14
			<input type="checkbox"/> SXT ⓘ	1991.09.03 – 2001.12.14
			<input type="checkbox"/> WBS ⓘ	1991.09.01 – 2001.12.14
<input type="checkbox"/> MWSPADP ⓘ		MtWilson ⓘ	<input type="checkbox"/> 60-ft SHG ⓘ	1915.08.10 – 1985.12.31
<input type="checkbox"/> NGDC ⓘ		GOES-12 ⓘ	<input type="checkbox"/> SXI-0 ⓘ	2001.09.10 →
<input type="checkbox"/> NSO ⓘ	<input type="checkbox"/>	Evans ⓘ	<input type="checkbox"/> spectroheliograph ⓘ	1996.02.05 – 1999.05.28
	<input type="checkbox"/>	GONG ⓘ	<input type="checkbox"/> Big Bear ⓘ	2005.04.11 →
			<input type="checkbox"/> Cerro Tololo ⓘ	2005.02.24 →
			<input type="checkbox"/> El Teide ⓘ	2005.02.25 →
			<input type="checkbox"/> Learmonth ⓘ	2005.02.25 →
			<input type="checkbox"/> MERGED GONG ⓘ	2001.07.22 →
			<input type="checkbox"/> Mauna Loa ⓘ	2005.04.11 →
	<input type="checkbox"/>	KPVT ⓘ	<input type="checkbox"/> 512-channel magnetograph ⓘ	1974.02.01 – 1993.04.10
			<input type="checkbox"/> spectromagnetograph ⓘ	1992.04.19 – 2003.09.21
	<input type="checkbox"/>	McMath ⓘ	<input type="checkbox"/> solar fts spectrometer ⓘ	1976.03.31 – 2003.12.31
	<input type="checkbox"/>	O-SPAN ⓘ	<input type="checkbox"/> O-SPAN ⓘ	2002.12.11 →
	<input type="checkbox"/>	SOLIS ⓘ	<input type="checkbox"/> vsm ⓘ	2004.01.02 →

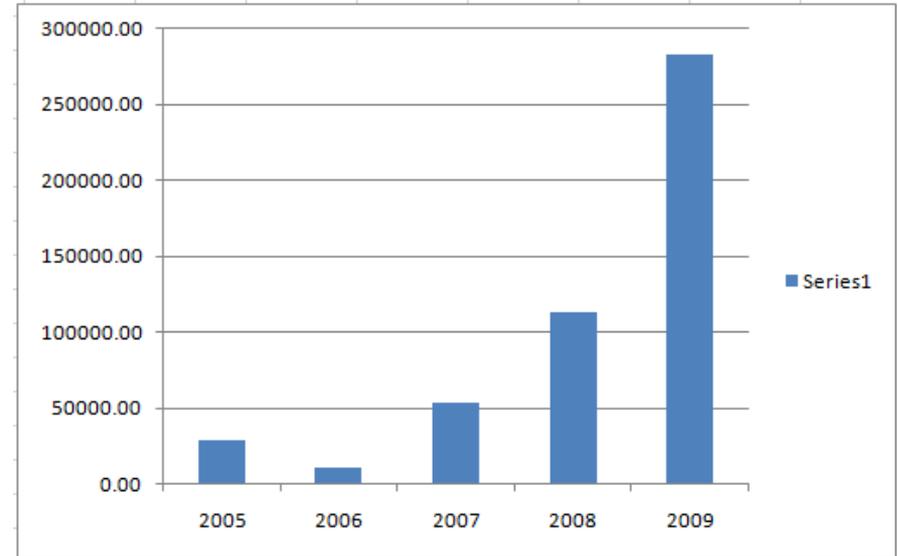
<input type="checkbox"/>	<a href="#">OBSPM</a>	<input type="checkbox"/>	<a href="#">Nancay</a>	<input type="checkbox"/>	<a href="#">Decametric Array</a>	2003.03.10 →
		<input type="checkbox"/>	<a href="#">OBSPM</a>	<input type="checkbox"/>	<a href="#">Radioheliograph</a>	1996.10.20 →
		<input type="checkbox"/>	<a href="#">Pic du Midi</a>	<input type="checkbox"/>	<a href="#">Meudon Spectroheliograph</a>	1995.12.01 →
<input type="checkbox"/>	<a href="#">OVRO</a>	<input type="checkbox"/>	<a href="#">OVRO</a>	<input type="checkbox"/>	<a href="#">Coronagraph</a>	1995.10.20 →
<input type="checkbox"/>	<a href="#">SAO</a>		<a href="#">Hinode</a>	<input type="checkbox"/>	<a href="#">OVSA</a>	2000.03.16 →
<input type="checkbox"/>	<a href="#">SDAC</a>	<input type="checkbox"/>	<a href="#">Hinode</a>	<input type="checkbox"/>	<a href="#">XRT</a>	2006.10.23 →
		<input type="checkbox"/>	<a href="#">SOHO</a>	<input type="checkbox"/>	<a href="#">EIS</a>	2006.10.23 →
				<input type="checkbox"/>	<a href="#">SOT</a>	2006.10.23 →
				<input type="checkbox"/>	<a href="#">CDS</a>	1996.01.19 →
				<input type="checkbox"/>	<a href="#">CELIAS</a>	1995.12.02 →
				<input type="checkbox"/>	<a href="#">COSTEP</a>	1995.12.07 – 2003.05.01
				<input type="checkbox"/>	<a href="#">EIT</a>	1996.01.01 →
				<input type="checkbox"/>	<a href="#">ERNE</a>	1996.05.08 – 2001.06.01
				<input type="checkbox"/>	<a href="#">GOLF</a>	1996.01.01 →
				<input type="checkbox"/>	<a href="#">LASCO</a>	1995.12.08 →
				<input type="checkbox"/>	<a href="#">MDI</a>	1996.05.01 – 2003.04.13
				<input type="checkbox"/>	<a href="#">SUMER</a>	1996.01.22 →
				<input type="checkbox"/>	<a href="#">SWAN</a>	1996.01.12 – 2003.01.01
				<input type="checkbox"/>	<a href="#">UVCS</a>	1996.01.20 →
				<input type="checkbox"/>	<a href="#">VIRGO</a>	1995.12.06 →
<input type="checkbox"/>	<a href="#">SFO</a>	<input type="checkbox"/>	<a href="#">TRACE</a>	<input type="checkbox"/>	<a href="#">TRACE</a>	1998.02.16 →
			<a href="#">SFO</a>	<input type="checkbox"/>	<a href="#">CFDT1</a>	1986.05.26 →
<input type="checkbox"/>	<a href="#">SHA</a>	<input type="checkbox"/>	<a href="#">GONG</a>	<input type="checkbox"/>	<a href="#">CFDT2</a>	1992.01.11 →
				<input type="checkbox"/>	<a href="#">Big Bear</a>	2001.03.14 →
				<input type="checkbox"/>	<a href="#">Cerro Tololo</a>	2001.04.20 →
				<input type="checkbox"/>	<a href="#">El Teide</a>	2001.07.30 →
				<input type="checkbox"/>	<a href="#">Learmonth</a>	2001.04.30 →
				<input type="checkbox"/>	<a href="#">Mauna Loa</a>	2001.06.16 →
				<input type="checkbox"/>	<a href="#">Udaipur</a>	2001.10.25 →
		<input type="checkbox"/>	<a href="#">JSPO</a>	<input type="checkbox"/>	<a href="#">MOTH</a>	2003.01.01 – 2003.01.20
		<input type="checkbox"/>	<a href="#">MtWilson</a>	<input type="checkbox"/>	<a href="#">MOF/60</a>	1996.05.01 →
		<input type="checkbox"/>	<a href="#">SOHO</a>	<input type="checkbox"/>	<a href="#">MDI</a>	1996.01.30 →
		<input type="checkbox"/>	<a href="#">TON</a>	<input type="checkbox"/>	<a href="#">Big Bear</a>	1996.06.01 – 1996.08.31
<input type="checkbox"/>	<a href="#">SSC</a>	<input type="checkbox"/>	<a href="#">STEREO_A</a>	<input type="checkbox"/>	<a href="#">Tenerife</a>	1996.06.03 – 1996.08.06
				<input type="checkbox"/>	<a href="#">IMPACT</a>	2006.10.01 →
				<input type="checkbox"/>	<a href="#">PLASTIC</a>	2006.10.01 →
				<input type="checkbox"/>	<a href="#">SECCHI</a>	2006.11.06 →
				<input type="checkbox"/>	<a href="#">SWAVES</a>	2006.10.27 →
		<input type="checkbox"/>	<a href="#">STEREO_B</a>	<input type="checkbox"/>	<a href="#">IMPACT</a>	2006.10.01 →
				<input type="checkbox"/>	<a href="#">PLASTIC</a>	2006.10.12 →
				<input type="checkbox"/>	<a href="#">SECCHI</a>	2006.11.07 →
				<input type="checkbox"/>	<a href="#">SWAVES</a>	2006.10.27 →

# Usage statistics

## Searches



## Files



These statistics are for one of three “Official” VSO instances

Statistics do not include usage through the IDL API, only those through the web interface

So far in 2010: 478 searches, 87,000 files, estimate 1400 searches, 250,000 files through VSO/NSO

# Next steps

- Finish SDO distributed data system
- Add spatial subset searches (currently restricted to full-disk images)
- Improve the user interface
- Continue adding data sets
- Add searches based on data computations?

# Interoperability

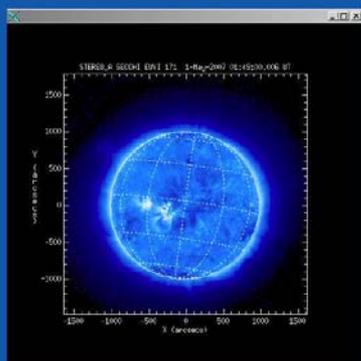
- Very important for VSO since no services are provided
- Examples:
  - IDL & SolarSoft
  - SDO/JSOC/HEK
  - Helioviewer
  - HelioScope, Aladin, TopCat
  - EGSO and successors
  - NASA Heliophysics VxOs

# IDL and SolarSoft

- IDL widely used in solar physics
- SolarSoft is a large library of procedures for analysis of NASA solar physics data
- VSO IDL client is an API written in IDL
- Uses SOAP/XML & HTTP/POST/GET to query VSO registry and retrieve matching datasets from providers
- `IDL> records=vso_search (tstart, tend, inst=inst, det=det, wave=wave)`
- `IDL> vso_get, records`
- Once the data is in IDL, have full range of analysis and plotting tools available

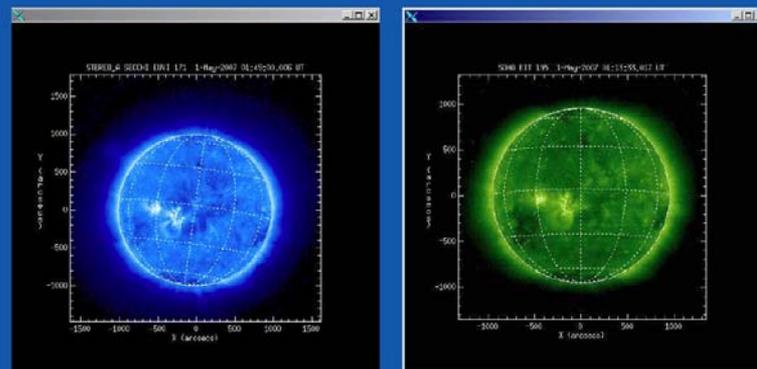
# An example using RHESSI Plotman

## Access STEREO Data



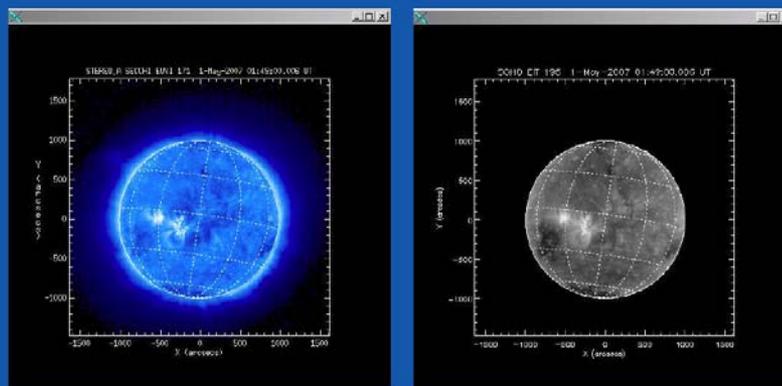
16

## Compare SOHO and STEREO



17

## Project SOHO to STEREO View



18

# SDO and VSO

- SDO (Solar Dynamics Observatory) is the new major NASA solar mission
- Largest data set in solar physics history (so far)
- Sustained 150 Mbs, 1.5 TB per day, expected total of 5 PB mission data set
- VSO is constructing a system of satellite sites to mirror part or the whole SDO archives
- Uses the Data Record Management System (NetDRMS) and Storage Unit Management System (SUMS) frameworks as the backbone
- Developed both by the SDO Joint Science Operations Center (JSOC) composed of Stanford University and Lockheed Martin
- Sites in Germany, Massachusetts, Arizona

# JSOC interface



jsoc.stanford.edu gives access to export series only, use jsoc2 for access to internal series

Series Select | Series Content | RecordSet Select | Values Display | Export Data | Test

### Information about selected series

Current Series is: aia\_test.lev1

PrimeKeys: T\_OBS, FSN

DBindex: T\_OBS\_index, FSN

Data NOT archived, online retention 60 days

Unit size: 1 record

Series Description AIA level 1 test

Release Notes [for aia\\_test](#)

This series contains no records

### 3. Select Records and Get Record Count

Enter RecordSet Specification here for keyword listings and for export.

Check box to show the QueryBuilder.

Request may take a while if the recordset is large (more than a few thousand records).

Record Limit  Optional, + for from start, - for from end.

Record Count:

- Check to Get Record Query.
- Check to Allow Huge Record Queries.
- Check to show full segment info
- Check to make local file links (only at JSOC).
- Prepare value table in 'show\_info' format in new window. (No \*psuedo\* keywords yet please) Segments fail

Select Keywords, Segments, and Links for table of values.

### 4. Select Keywords

cparams\_sg000  
image\_lev1\_bzero  
image\_lev1\_bscale  
cparams\_sg001  
bad\_pixel\_bzero  
bad\_pixel\_bscale

### 5. Select Segments

image\_lev1  
bad\_pixel  
spikes

### 6. Select Links

# Heliophysics Event Knowledgebase (HEK)

## The Sun 2010-5-17

- 1 Day +

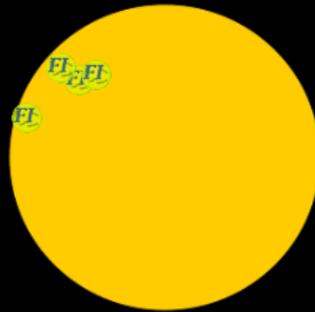
Last Updated 17-May-10 13:14:11.130 UT

"What's this?"

### Events



Disk Carrington Map



<< 2010-05-17T00:00:00, >>  
2010-05-18T00:00:00



[about / contacts](#)

Search results ([export](#))

- 1.FI: Filament
- 2.FI: Filament
- 3.FI: Filament
- 4.FI: Filament

### iSolSearch (v1.1)

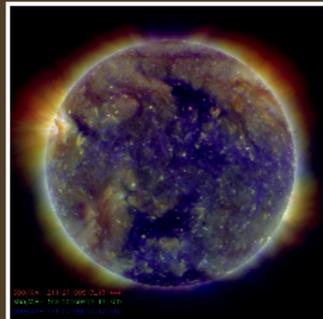
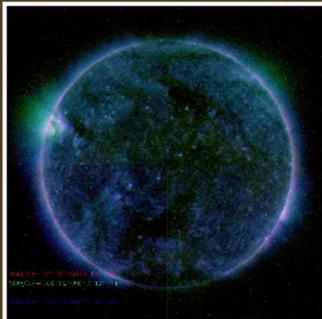
Welcome to iSolSearch – a window into the Heliophysics Events Knowledgebase (HEK). iSolSearch and HEK are designed to guide heliophysics researchers to relevant solar data, and primarily to that acquired by the Solar Dynamics Observatory (SDO) with the Atmospheric Imaging Assembly (AIA) and the Helioseismic and Magnetic Imager (HMI).

- [Users Guide](#)
- [QuickTime User Video](#)
- [SolarSoft IDL](#)
- [API Documentation](#)

other links...

- [Heliophysics Events Knowledge Base](#)

### e Channels



Not yet available

# Helioviewer

 Helioviewer.org

**Time**

Date:

Time:

Time-step:

**Images** [Add]

▸ EIT 304 2010/01/01 01:19:34 |

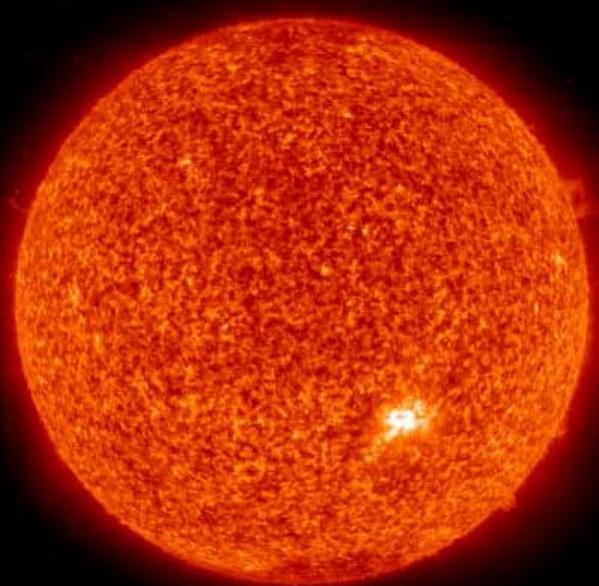
**Features/Events** [Add]

▸ LASCO CME List C |

▸ GOES X-Ray G |

▸ NOAA Active Regions AR |

**center** 🔗 Link ⬇️ JHelioviewer 🗄️



Help About Usage Tips Wiki API Contact Report Bug

# HeliScope, Aladin, TopCat via Plastic

The image displays two software windows from the AstroGrid Workbench. The main window is HeliScope, showing search results for the period 2006-01-19T12:00:00 to 2006-01-19T18:00:00. The search results are organized into a radial tree structure with a central node 'Time Series/Images' and several branches representing different data sources: Cluster spacecraft dataset, Polar spacecraft dataset, Ulysses spacecraft dataset, RHESSI (Reuven Ramaty High), SOHO Coronal Diagnostic, SOHO LASCO (Large Angle), LASCO, SOHO EIT (Extreme ultraviolet), EIT, TRACE (Transition Region And), and ACE spacecraft dataset. The ACE spacecraft dataset branch is further detailed with a list of specific data files and their timestamps, including '2006-01-19T18:00:00' which is highlighted in yellow.

The second window is Aladin v3.7 multiview, displaying two astronomical images. The left image is labeled 'PLASTIC-1' and shows a bright circular object. The right image is labeled 'PLASTIC' and shows a grayscale image of a celestial body, likely the Moon, with a pink crosshair and a red arrow pointing to a specific feature. The Aladin window includes a toolbar with various tools like 'select', 'dist', 'draw', 'tag', 'text', 'filter', 'rgb', and 'assoc', and a status bar indicating '2 planes, 2 views, 2Mb'.

# EGSO interface



## European Grid of Solar Observations

[Introduction](#) | [Help](#) | [Tools](#)

**EGSO GUI - BUILD QUERY**

**New query** **Build query** **List of observations** **List of files**

**QUERY TYPE**

Standard Query

**SEARCH PARAMETERS**

Remote Observable Entity

Insitu Observable Entity

**RESULT PARAMETERS**

Select Plots

**DATE/TIME**

Start Date: 2002-07-15 00:00:00 ...

Format: 2003-12-31 23:59:59

End Date: 2002-07-16 00:00:00 ...

Format: 2003-12-31 23:59:59

**EVENT**

None

GEV GOES Event List

NOAA Proton Events

**REMOTE OBSERVABLE ENTITY**

Select all Photons

Photons:  Gamma Rays

XRays:  SXR  HXR

Ultra Violet:  EUV  UV

Visible (inc. H alpha and 10830A)

Infrared

Microwaves

Radio Waves

**INSITU OBSERVABLE ENTITY**

Select all Particles

Particles:  Energetic

Charged

Neutral

Select all Fields

Fields:  Electrical

Magnetic

Gravitational

**SELECT PLOTS**

none

GOES XRay Plot

GOES Proton Plot

**New query** **Build query** **List of observations** **List of files**

# NASA Heliophysics VxOs

- Next talk by Aaron Roberts
- Several different flavors of virtual observatories for specific disciplines in heliophysics
- A central data model and framework to glue them all together

# Lessons learned

- Basic design also used by NVO, PDS, so probably the optimal solution
- Best Decisions
- What had to be revised
- User uptake

# Best decisions (in hindsight)

- Chose SOAP over various competing methodologies
- Were forced to KISS (“small box”) because of very constrained budget
- Few non-data services, no public outreach features
- Supported by SolarSoft for a good fraction of community’s needs
- Continually went to community for input
  - AAS, SPD, AGU meetings/BoF’s
  - Beta testing and feedback

# What we had to change or add

- Registry
- Replication of registry
- Data model (still being expanded)
- Writing SOAP servers for some data providers
- Providing proxies for database-less providers
- Even some hardware for one data provider

# User uptake

- Resistance by elders (“I don’t like it”) is hard to overcome
  - Keep the old access methods, i.e. direct FTP
- Younger researchers adopt quickly
- Interfaces are hard to do well
  - Too much box clicking is bad (VSO is still guilty)
  - More graphical systems highly dependent on browsers

# Conclusion

- The simple design has resulted in a robust system that can be maintained with limited funding
- Many other groups are developing data services for the VSO
- VSO is increasingly being used by the heliophysics community