F. The Virtual Solar Observatory

In addition to the "classic" data center roles described above, the SDAC has provided leadership and software development for the Virtual Solar Observatory. We proposed a study and implementation plan for the VSO in our FY2001 Senior Review proposal, and have carried out that plan in partnership with groups we fund at Stanford and Montana State Universities, Southwest Research Institute, and the National Solar Observatory. Some milestones of VSO development are listed in Table II-2.

Milestone	Date
Proposed as "optimal" delta to Senior Review	2001 April
Study report (design document) released for community comment	2002 November
VSO 1.0 release: initial data model (dictionary); Stanford, Montana State, NSO, and SDAC data providers	2003 December (AGU)
VSO 1.1 release: searches on event catalogs added; added HAO Mauna Loa, OVRO	2004 June (AAS/SPD)
VSO 1.2 release: improved "shopping cart;" thumbnail and browse images; original data providers, plus RHESSI, TRACE, GOES SXI, H Alpha Network, Meudon, SOLIS	2005 May (AGU)
Added Mt. Wilson Ca II spectroheliogram (starting in 1915), San Fernando Observatory, O-SPAN data providers	2005 - 2006

Table II-2. VSO milestones.

The initial VSO study immediately recognized that in the time and with the resources available, we could not solve all problems for all solar physics data identification and access. We did conclude, however, that we could provide all the services delimited by the "small box" of Figure 1, while imposing the minimum of requirements on the data providers. All they would have to establish would be a simple Web service based on the Simple Object Access Protocol (SOAP), with which one or more instances of the VSO server could communicate. (In practice, there are VSO servers at Stanford, NSO, and the SDAC; it is also possible for any user or facility to run their own instance, as we have done on laptops while demonstrating the VSO at several scientific meetings.)

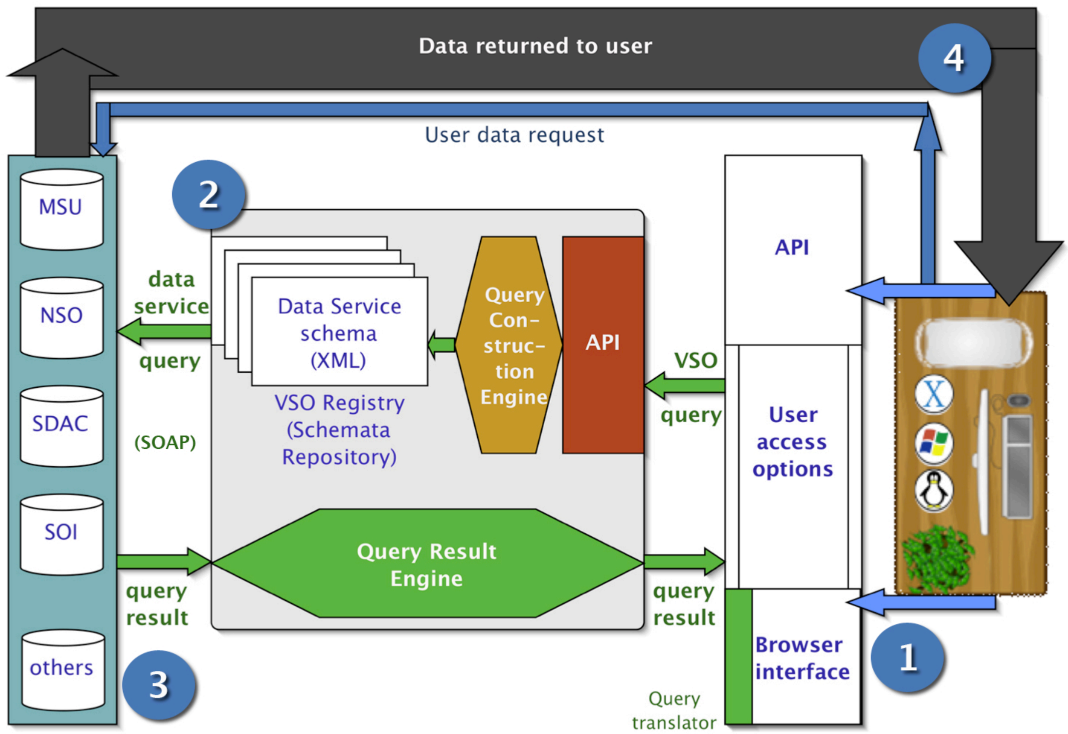

Figure 1. The VSO "small box:" in addition to an API, the VSO consists only of a query construction engine, a registry of XML schema describing each data providers' searchable holdings, and a method of communicating the queries to the appropriate providers and the results to the user. When used with the Web browser interface, the VSO returns its results in the form of a unique "shopping cart," which can be cited and revisited. Note that the VSO itself delivers no data, only links to data objects on the data providers' servers.

Readers of this proposal are invited to exercise the VSO search capabilities at: http://virtualsolar.org/ .

The VSO appears to be well on its way to being a standard tool used by solar physicists and those in related fields to identify and access research-quality data. Figure 2 shows the growth in VSO usage over the last three quarters.

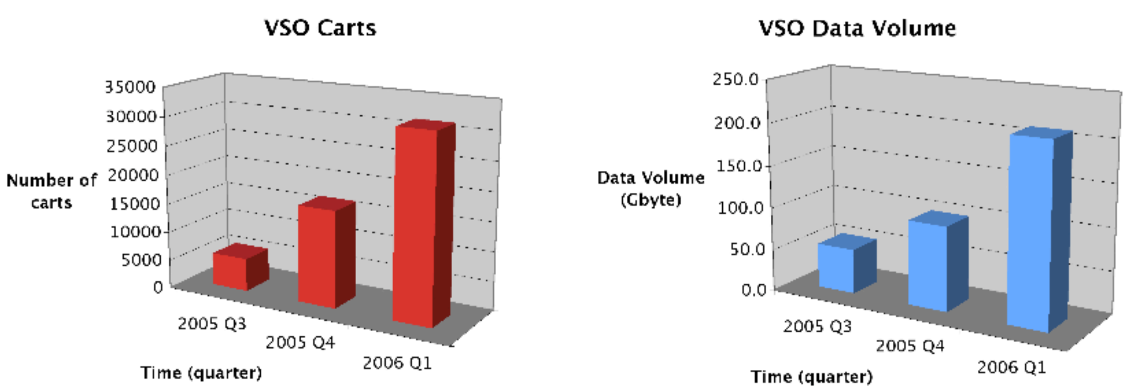

Figure 2. VSO usage trends, calendar year 2005 Q3 – 2006 Q1: *(left)* individual VSO data "shopping" carts, the most direct measure of the number of distinct searches carried out by

live users of the VSO Web interface; *(right)* total volume of data *identified* (and links provided) by VSO Web interface searches. We start with 2005 Q3 to prevent "contamination" by the large number of demonstration carts and data links amassed at our booth at the 2005 New Orleans AGU/Solar Physics Division Joint Assembly. Note that API accesses of the VSO are *not* included in these statistics.

It should be noted that Figure 2 represents only live, human interaction with the VSO *via* our Web interface. Machine access, using the VSO itself as a Web service to identify and provide links to solar physics data, is as yet uncounted, though we plan to implement such counting in the next year. An example of this type of usage of the VSO is the UK AstroGrid's HelioScope data discovery and delivery interface (see Figure 3), written by Sylvia Dallia of the University of Manchester.

The VSO was envisioned in our FY2001 proposal as a way to replace all the centralized aspects of items A – C, above, while maintaining expertise in the mission and the data at the institutions where the instruments were built or operated. Indeed, we argued that the SDAC should "wither away," and if the community and NASA management felt that the SDAC no longer added significant value, it should be defunded. We have to confess that we have failed at replacing the SDAC with a completely distributed solar physics data system based on the VSO, but we believe the work breakdown in the next Section demonstrates that we have reduced the SDAC to the bare minimum (or perhaps less) that is useful to the community in achieving its part in new system science of heliophysics, as summarized in Strategic Goal 3B of the NASA 2006 Strategic Plan: "Understand the Sun and its effects on Earth and the solar system."

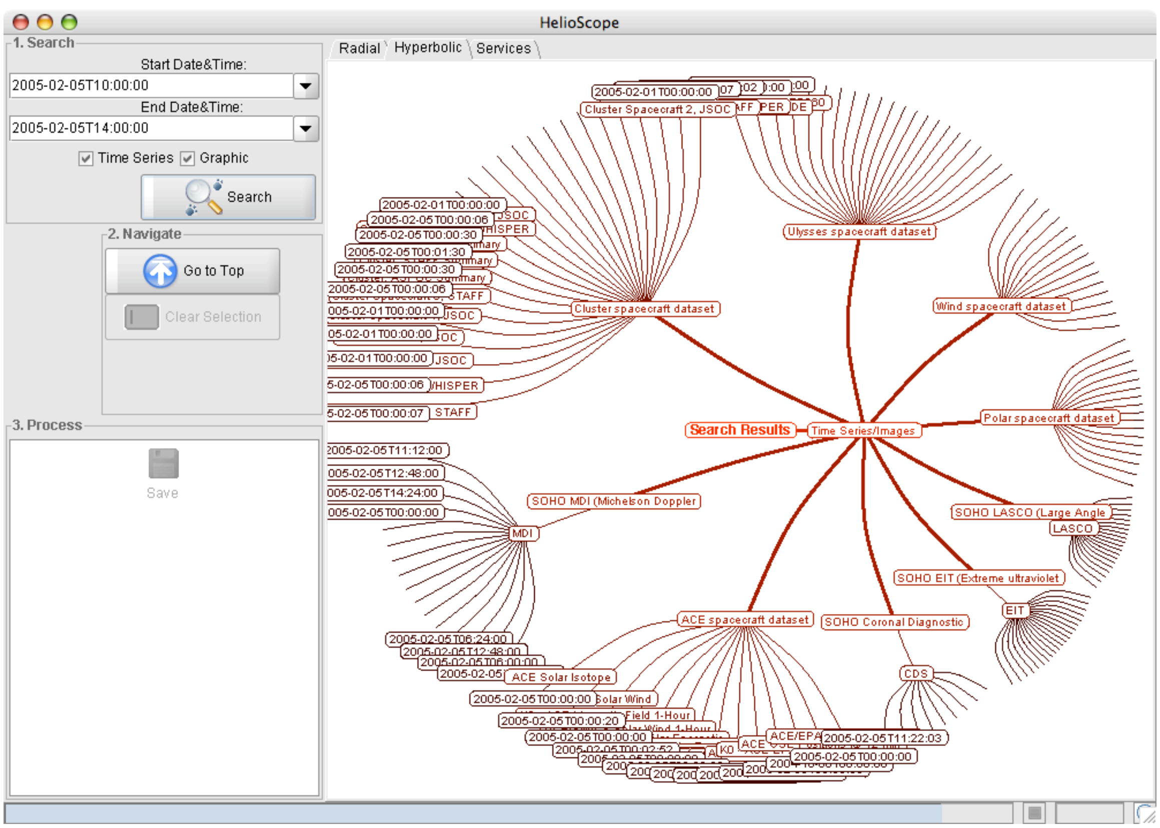

Figure 3. One representation of the results of a search for heliophysical data with the AstroGrid HelioScope tool. Heliospheric and magnetospheric data have been identified *via* the CDAWeb; solar physics data (*SOHO* and [not shown] TRACE) *via* the SDAC's Web service API. Written in Java, HelioScope is a front end for a Grid-enabled data search and manipulation facility.

H. VSO Baseline Requirements

Since the VSO now forms such a substantial fraction of our baseline effort, we break the VSO work down one level further to make it clear where our procurement resources are currently directed. In accordance with our FY2001 proposal, we regard the ***baseline*** VSO tasks in FY07 and beyond as being limited to maintenance and development of only those features necessary to accommodate the inclusion of new data services, such as the STEREO, Solar-B, and SDO missions.

1.3.1 Incorporate access to new NASA mission data (at a minimum, STEREO, Solar-B, SDO)

1.3.2 Incorporate access to additional groundbased data sources

1.3.3 Assist data providers with constructing services that will be VSO-accessible

1.3.4 Assist providers of small but critical data sets with small hardware assistance if that is all that stands between them and the VSO (*e.g.* recently digitized Mt. Wilson Ca II imagery beginning in 1915)

1.3.5 Add more catalogs/event lists

1.3.6 Add the capability to search by data *cadence*

1.3.7 Improve performance (important as the cadence grows, *e.g.* in SDO)

1.3.8 Insure that internals provide scalability

1.3.9 Add higher-level helioseismology data products

1.3.10 Assure API usability

1.3.11 Respond to user trouble reports

1.3.12 Management (project management, communicating status and capabilities to the scientific community, NASA procurement)

I. Optimal VSO Requirements

The VSO implementation team strongly believes that the real power and scientific utility of the VSO as a tool for heliophysics *system* science can only be achieved through a more substantial but still modest-cost augmentation of its capabilities. Those augmentations are listed in the following breakdown:

1. Develop the full potential of the VSO

    1.1 Provide the ability to join searches on multiple catalogs/event lists and data sources

1.2 Provide ability to search on spatial regions/coordinate ranges of interest as well as time, &c.

1.3 Improve robustness/availability of VSO server instances

1.3.1 software recognition of when specific servers are unavailable

1.3.2 hardware for redundancy, failover

1.4 Work with other VO's on building a heliophysics meta-VO

1.4.1 Metadata translations

1.4.2 API change control, backwards compatibility

1.4.3 Community planning efforts (*e.g.* NASA Heliophysics)

1.5 Work toward making VSO compatible with semantic Web

1.5.1 Ontologies, Resource Description Framework (RDF)

1.5.2 Community efforts (*e.g.* AGU)

1.6 Graphic User Interface (GUI) improvements

1.7 Connections to applications packages (*e.g.* CoSEC)

1.8 Establish database of carts, publications, and other user annotations

1.9 Enable Cart ID's for API calls (to allow citation, statistics)

Management (project management, communications with larger community and NASA management

VSO Work breakdown				FY07		FY08		FY09		FY10		Notes
				Labor (FTE)	Hardware ($K)	Labor (FTE)	Hardware ($)	Labor (FTE)	Hardware ($)	Labor (FTE)	Hardware ($)	
I. VSO baseline requirements												
	1.3.1		Incorporate access to new NASA mission data (at a minimum, STEREO, Solar-B, SDO)	0.38		0.30		0.19		0.15		A
	1.3.2		Incorporate access to additional groundbased data sources	0.30		0.30		0.25		0.25		
	1.3.3		Assist data providers with constructing services that will be VSO-accessible	0.08		0.08		0.08		0.08		
	1.3.4		Assist providers of small but critical data sets with small hardware assistance if that is all that stands between them and the VSO (e.g. Mt. Wilson)	0.02		0.02		0.02		0.02		
	1.3.5		Add more catalogs/event lists	0.15		0.15		0.07		0.04		
	1.3.6		Add cadence searching	0.06		0.06		0.25		0.12		
	1.3.7		Improve performance	0.20		0.05						
	1.3.8		Insure that internals provide scalability	0.00		0.25		0.05		0.05		
	1.3.9		Add higher-level helioseismology data products	0.05		0.05						
	1.3.10		Assure API usability	0.25		0.25		0.13		0.10		
	1.3.11		Responding to user trouble reports	0.23		0.23		0.23		0.23		
	1.3.12		Management	0.10		0.10		0.10		0.10		C
	Totals			1.82		1.84		1.37		1.14		
	Notes											
A			Some work under way already for STEREO and SDO, paid for by SDAC and their Phase C/D budgets; FY10 figure is a placeholder for TBD mission(s)									
C			Civil service labor									

Our "optimal" budget consists of our baseline budget plus the Optimal VSO work described in Section II.H.

VSO Work breakdown				FY07		FY08		FY09		FY10		Notes
II. New work ("optimal" budget)												
7			Develop the full potential of the VSO									
	7.1		Provide ability to join searches on multiple catalogs/event lists and data sources	0.40		0.30		0.15				
	7.2		Provide ability to search on regions/coordinate ranges of interest as well as time, &c.	0.19		0.12		0.06				
	7.3		Improve robustness/availability (software recognition									
		7.3.1	software recognition of when specific servers are unavailable	0.07		0.03		0.03		0.02		
		7.3.2	hardware for redundancy, failover	0.01	9.0	0.01		0.01				B
	7.4		Work with other VO's on building a heliophysics meta-VO									
		7.4.1	Metadata translations	0.20		0.04		0.04		0.04		
		7.4.2	API change control, backwards compatibility			0.08		0.06				
		7.4.3	Community planning efforts	0.10		0.10		0.10		0.10		
	7.5		Work toward making VSO compatible with semantic Web									
		7.5.1	ontologies, RDF	0.15		0.35		0.15		0.15		
		7.5.2	community efforts (e.g. AGU)	0.10		0.10		0.10		0.10		
	7.6		GUI improvements	0.10		0.10		0.10		0.10		
	7.7		Connections to applications packages (*e.g.* CoSEC)	0.10		0.13		0.20		0.20		
	7.8		Establish database of carts, publications, and other annotations	0.28		0.05		0.04		0.04		
	7.9		Cart ID for API calls			0.29		0.20				
	7.10		Management	0.10		0.10		0.10		0.10		C
	Totals			1.70	9.0	1.70		1.24		0.75		
	Notes											
B			$4K at SDAC, $5K at SHA for additional server hardware									
C			Civil service labor									

The detailed budget is given in the separate spreadsheet document. We discuss only highlights and impacts.

[snip]

"Optimal" budget: fulfilling the promise of the VSO. Having taken the plunge in funding virtual observatories in Heliophysics through the ROSES VO line, the LWS TR&T program, and this review, the Heliophysics Division must decide whether it will reap the full benefit of these efforts and fund work to integrate the various discipline VO's into a usable meta-VO. The next logical step is to examine whether the VO's can become part of the semantic Web (http://www.w3.org/2001/sw/) that will allow Internet users to tap the "undiscovered country" of Web services not represented in the flat-HTML pages to which current search engines (*e.g.* Google) are largely limited.

The optimal VSO budget also asks for funding to enable searches based on joins of multiple catalogs (flares, CME's, active regions, &c.) and the metadata describing the observational data more directly. Thus a user could, for instance, search for EUV data corresponding to white-light CME's in certain latitude regions with underlying active regions in a user-specified range of magnetic flux, with eruptive prominences in a user-specified plane-of-the-sky speed range. We believe that a fully joined search capability of this kind is critical to exploiting the observations to be made with STEREO, Solar-B, and SDO, so our proposed "optimal" budget is front-loaded to implement these capabilities as soon as possible.

SDAC Bibiliography

Bogart, R. 2004, "The Virtual Solar Observatory," B.A.A.S. 36, Abstract 52.04, 754

Bogart, R. S., Davey, A., Dimitoglou, G., Gurman, J. B., Hill, F., Martens, P. C., Tian, K. Q., and Wampler, S. 2003, "Data Integration in the Virtual Solar Observatory," American Geophysical Union, Fall Meeting 2003, abstract #SH42A-0503

Bogart, R.S., Tian, K.Q., Davey, A., Dimitoglou, G., Gurman, J., Hill, F., Hourclé, J., Martens, P.C., Suárez-Sola, I., Tian, K. Q., Wampler, S., and Yoshimura, K. 2005, "Building A Virtual Solar Observatory: Lessons Learned," Eos Trans. AGU, 86(18), Jt. Assem. Suppl., Abstract SH51B-05, Pg. JA460

Davey, A. R., Bogart, R. S., Dimitoglou, G., Gurman, J. B., Hill, F., Martens, P. C., Tian, K. Q., and Wampler, S. 2003, "First Steps Towards a VSO," American Astronomical Society, SPD meeting #34, #03.11

Davey, A. R., Bogart, R. S., Gurman, J. B., Hill, F., Hourcle, J., Martens, P. C., Suarez-Sola, I., Tian, K. Q., and Yoshimura, K. 2004, "Implementation of the Virtual Solar Observatory," American Astronomical Society Meeting 204, 2004AAS...204.7002D

Davey, A. R., Wills-Davey, M., Gurman, J. B., Bogart, R., Tian, K. Q., Martens, P., Yoshimura, K., Hill, F., Suarez-Sola, I., Hourclé, J., Dimitoglou, G., and Wampler, S. 2004, "Doing Science with the VSO: Signatures of CME Initiation," American Geophysical Union, Fall Meeting, 2004AGUFMSH21B0414D

Davey, A., Wills-Davey, M., Gurman, J., Bogart, R., Dimitoglou, G., Hill, F., Hourcle, J., Martens, P., Suarez Sola, I., Tian, K., Wampler, S., and Yoshimura, K. 2005, "CME Population Distributions: Science Facilitated by the VSO," American Geophysical Union, Spring Meeting, 2005AGUSMSP23A..09D

Gurman, J.B. 2001, "Toward a Virtual Solar Observatory: Starting Before the Petabytes Fall," American Astronomical Society, 199th AAS Meeting, #112.03, Bulletin of the American Astronomical Society, Vol. 33, 1478

Gurman, J.B. 2002, "Toward a Virtual Solar Observatory," in Proceedings of the SOHO 11 Symposium on From Solar Min to Max: Half a Solar Cycle with SOHO, 11-15 March 2002, Davos, Switzerland. A symposium dedicated to Roger M. Bonnet. Edited by A. Wilson, ESA SP-508, Noordwijk: ESA Publications Division, ISBN 92-9092-818-2, 2002, p. 525

Gurman, J.B. 2004, "Virtual Observatories: Are We Virtually There Yet?" (Parker Lecture), American Astronomical Society Meeting 204, #52.01

Gurman, J. B., Bogart, R., Hill, F., and Martens, P. 2002, "Building a Virtual Solar Observatory: I Look Around and There's a Petabyte Following Me," Bulletin of the American Astronomical Society, Vol. 34, p.740

Gurman, J., Bogart, R. S., Davey, A.R., Dimitoglou, G., Hill, F., Hourclé, J., Martens, P.C., Suárez-Sola, I., Tian, K. Q., Wampler, S., and Yoshimura, K. 2005, "The Virtual Solar Observatory: Still A Small Box," Eos Trans. AGU, 86(18), Jt. Assem. Suppl., Abstract SH43B-03, Pg. JA456-JA457

Gurman, J. B. and Davey, A. R. 2004, "Future Mission Data Environment: Virtualizing Access to Solar Physics Data," in Proceedings of the SOHO 15 Workshop - Coronal Heating. 6-9 September 2004, St. Andrews, Scotland, UK (ESA SP-575). Editors: R.W. Walsh, J. Ireland, D. Danesy, B. Fleck. Paris: European Space Agency, p.583

Gurman, J. B., Dimitoglou, G., Bogart, R. S., Tian, K. Q., Hill, F., Wampler, S., Martens, P. C., and Davey, A. 2002, "The Right Amount of Glue: Technologies and Standards Relevant to a Future

Solar-Terrestrial Data Environment," American Geophysical Union Fall Meeting, 2002AGUFMSH52C..03G

Gurman, J. B., Dimitoglou, G., Hill, F., Wampler, S., Bogart, R. S., Tian, K., Martens, P., and Davey, A. 2003, "The Virtual Solar Observatory: The Best-Laid Schema," American Astronomical Society, SPD meeting #34, 2003SPD....34.0203G

Gurman, J. B., Dimitoglou, G., Bogart, R. S., Davey, A. R., Hill, F., Martens, P. C., Tian, K. Q., and Wampler, S. 2003, "Virtual Solar Observatory Distributed Query Construction," Eos Trans. AGU, 84(46), Fall Meeting Suppl., Abstract U22A-0011, Pg. F11

Hill, F., Bogart, R.S., Davey, A., Dimitoglou, G., Gurman, J.B., Hourclé, J.A., Martens, P.C.,

Hill, F., 2004, "What's Inside the Small (VSO) Box," presentation at Space and Solar Physics Virtual Observatories Workshop, Greenbelt MD

Hill, F., Bogart, R. S., Davey, A., Dimitoglou, G., Gurman, J. B., Martens, P. C., Tian, K., and Wampler, S. 2003, "The Virtual Solar Observatory – The Why, What, How and Where," Eos Trans. AGU, 84(46), Fall Meeting Suppl., Abstract U22A-0012, Pg. F11

Hourclé, J.A., Bogart, R., Davey, A., Gurman, J. B., Hill, F., Martens, P., Suárez-Sola, I., Tian, K., and Yoshimura, K. 2004, "STEREO In The Virtual Solar Observatory Context," Eos Trans. AGU, 85(47), Fall Meeting Suppl., Abstract SH21B-0413, Pg. F1484

Hourclé, J., Dimitoglou, G., Suárez -Sola, I., Tian, K., and Davey, A. 2004, "Scalability of the VSO," B.A.A.S. 36, Abstract 70.04, 796

Hourclé, J., Davey, A., Suárez -Sola, I., Tian, K., and Yoshimura, K., 2004, "Data Transfer Negotiation with the Virtual Solar Observatory," presentation at Space and Solar Physics Virtual Observatories Workshop, Greenbelt MD

Martens, P. C. H., Gurman, J. B., Hill, F., Bogart, R. S., Davey, A., Dimitoglou, G., Tian, K., and Wampler, S. 2002, "Strawman Concept for A Virtual Solar Observatory," American Astronomical Society Meeting 201, 2002AAS...20114103M

Suárez-Sola, I., Tian, K., Wampler, S., and Yoshimura, K, 2004, "The Virtual Solar Observatory: status and initial operational experience," Ground-based Telescopes. Edited by Oschmann, Jacobus M., Jr. Proceedings of the SPIE, 5493, p. 163-169

Suárez -Sola, I., Bogart, R., Davey, A., Dimitoglou, G., Gurman, J., Hill, F., Hourclé, J., Martens, P., Tian, K., Wampler, S., and Yoshimura, K., 2004, "Virtual Solar Observatory Applications," presentation at Space and Solar Physics Virtual Observatories Workshop, Greenbelt MD

Tian, K. Q., Bogart, R. S., Davey, A., Dimitoglou, G., Gurman, J. B., Hill, F., Martens, P. C., and Wampler, S. 2003, "Data Integration Using SOAP in the VSO," American Astronomical Society, SPD meeting #34, #03.12