

VSO: Technical Status, Issues, and Lessons Learned

Joe Gurman and Joe Hourclé
NASA Goddard Space Flight Center

Igor Suàrez-Sola
National Solar Observatory

VSO Status and Priorities (I)

- Emphasis in last ~ 6 months has been on integration of new data providers
 - New missions
 - STEREO: largely complete, also use VSO engine as back-end on STEREO Science Center Website
 - *Hinode*: work just beginning as two of three instruments data are reprocessed

VSO Status and Priorities (II)

- Priorities for next 6 months:
 - Cart API
 - So that other UIs can save / load carts
 - Catalog API / Catalogs as a Service
 - Generic ways to describe / store catalogs
 - Registry API
 - To allow UIs to ask about the general holdings of VSO.
 - Customizable Web UI
 - To allow alternate look & feel
 - Allow data providers to use VSO framework but still pass archive-specific parameters

VSO Status and Priorities (III)

- Priorities for next 6 months (*continued*):
 - New Parameters / Updated Data Model
 - Dynamic Metadata Return
 - Requesting fields other than the standard VSO return structure be returned
 - Must be supported by the Data Provider being asked
 - Data Provider Customizable Parameters
 - Ways for the Data Provider to note useful parameters used by their system, that aren't in the data model.

VSO Issues and Lessons Learned

Data Transfer Negotiation

- The Problem:
 - There is no one 'best' data transfer method
- VSO's solution: GetData
 - Initial 'Query' only returns metadata with Foreign Keys
 - Once users determine which data they want, they can request it with their preferred transfer methods
- For the Future:
 - This could be extended to handle local resolution
 - Or to negotiate file formats

IDL SOAP Client

- *The status quo:*
 - IDL is the preferred environment of the community
- VSO's solution:
 - A SOAP client in IDL
- For the Future:
 - Rewrite to deal with larger data structures
 - Need to test against other SOAP servers

Summary Metadata: Dealing with Large Datasets

- Summary metadata:
 - unique provide ID
 - e.g. GONG magnetogram Big Bear 10 min avg. id = 11005
 - total product count
 - total size
 - a URL identifier for the actual summary data
 - e.g. <http://vso.nso.edu/cgi/stream?pptid=11005;date=20070502>
- Advantages:
 - Lower overhead transaction
 - VSO Cart contains unique identifier for later re-query and publication completeness
 - Faster and easier access through the summary URL

VSO GUI and Cart

- Moving the CGI toward a pluggable interface that allows customization
 - By site or user
- Cart
 - Editing functionality
 - Allowing a Wiki-like info field
 - Downloadable movies
 - Publication linking

Lessons Learned

- Design for easy maintenance
 - You won't have full funding forever
 - Keep the central logic as simple as possible
 - Handle special cases as wrappers around the problem component
 - Your data model will change
 - Be careful about outthinking yourself