

DAP, SODA, Datalink status

F.Bonnarel (CDS)

reporting DAL working group
discussions



SIAP2/SODA evolution

Recent DAL running meeting

- https://wiki.ivoa.net/twiki/bin/view/IVOA/IvoaDAL_RunningMeetings#RM14

Running Meeting #14 - Wed. 28 September, 1pm UTC

agenda:

- Dataset Access Protocol (or SIA-next or [SODA-next](#)) [-slides-](#)

meeting notes: [Etherpad link](#)

R U N N I N G



DAP : main point of agreement

- SIA2.0 scope has to be extended to generic dataset protocol : DAP (or SDAP, or DsSAP)
 - Extend parameter interface to all data product types datasets
 - This new protocol will be a way to have for the same price SIA2.1, SSA2.0 and TimeSeriesAccess 1.0 + eventlist, visibilities, etc...
 - Name issue : these protocols are no more simple. But isn't DAP ambiguous (ObsTAP is a Dataset acces protocol) → p(arameter)DAP ?
- New DAP github repository : add your issues, comments, proposals there !!!



Some more critical points in « DAP »

- Several proposals (MOC parameter, case insensitivity, mandatory autodescription,...) seem to be accepted. See discussion of the running meeting quoted above.
- 1st critical point : allow wild carding in string-type parameters ?
 - If we use * we may have confusion for some strings. Character encoding may add some complexity
 - Multiple query seems difficult to replace wildcarding for obs_publisher_id eg.



Some more critical points in « DAP »

- 2nd critical point : one shot discovery of cutouts/rebinned-projected images
 - Very important to extract data from « mosaics » including « HiPS » in one shot (batch or massive repeated queries)
 - CDS prototype based on HiPStoFITS
 - Solution 1 : use soda-like (or better SODA) url in the access_reference FIELD → how to access the full dataset ?
 - Solution 2 : use service descriptor in query response to build the soda url → more work for client



SODA next

- Errata, MOC parameter, Rebin/reproject, pixel cutout, ... ;
- Choose dataproduct_type of retrieved data. Eg :
 - Extract spectrum or image from cube
 - Extract TimeSeries from catalog.
- Multiple Id / Parameter crossing
 - POST + tar file with results ?
(Vialectea test implementation)
- Metadata extraction (fits header, ObsCore extensions, provenance?) with new parameter ?



Towards DataLink 1.1 recommendation

- DataLink 1.1 working draft since one year. Kept idle since then
- Urgent to get some implementations (see below)
- Reminding the main changes



Towards DataLink 1.1 recommendation

- Extend the scope of DataLink when the starting is not a dataset (eg a source or an event in a catalog)
- INFO tag at the beginning of the response
- Allow fragmented URL references to the link
- Service descriptors enhancement
 - Names
 - DESCRIPTION
 - Nesting
 - content_type parameter
 - Better description of input parameters



Towards DataLink 1.1 recommendation

- Removing the getavailability section (get availability is no more a standard in VOSI)
- Announce authorization/authentication in advance by optional FIELDS
- Introduce the content_qualifier FIELD to qualify the nature of the retrieved resource :
 - today used to give the dataproduct-type.
 - Could not be done by content_type
- Service autodescription better explained



Towards DataLink 1.1 recommendation

- Forgotten :
 - Row index in links response to identify similar links for different #this → local semantics FIELD has been proposed, but is not yet in the draft
- Unsolved , delayed
 - Moving Service descriptors to VOTable spec ?
 - URL templating in service descriptor
 - Use of PQL in service descriptor
 - optional/mandatory distinction in service descriptor
PARAMETERs
 - Wrong UCD for ID FIELD in links response



Towards DataLink 1.1 recommendation

- Implementation
 - A lot of implementations of 1.0 now (this interop)
 - DACHS has non-datasets DataLink capabilities and local semantics
 - CADC to implement features in servers next November
 - TopCAT manages most of changes (eg service descriptors) already
 - Aladin plans to manage that before next interop
 - More implementations ? Needed !!!

