

DAL WG Status

D.Tody, June 2004

Discussions in Boston focused on all aspects of the Simple Spectral Access interface (SSA), various enhancements and upgrades to SIA, and the longer term strategy for integration of query language capabilities (ADQL) into the DAL services.

DAL Services

Simple Spectral Access (SSA)

The scope of SSA is spectrophotometric data in tabular form: 1D spectra, time series, and SEDs. It does not include spectral image cubes, longslit spectra, or synoptic imagery, which will be handled by the image service since they are gridded data. Much of the work being done on the SSA data model will however carry over to image data.

The components of SSA are:

- The SSA data model.
- The query interface, including query and response table.
- Data access, including standard representations of spectral and time series data in various formats.

All aspects of SSA were discussed in Boston and working draft versions of the interface documents are in preparation. Much of this work is being done in collaboration with the other working groups, e.g., the SSA data model is a joint effort of the DM and DAL groups, we are working with the VOTable WG on data representation issues, and with the VOQL WG on queries.

The VO community appears to be primed and ready to go with service implementations and client applications once the interface specification is ready. A quick survey of those present in Boston found 8 sites planning to implement spectrum or SED services for atlas (pre-generated) data, 2-3 sites which would implement services to generate spectral data on the fly, and 2-3 sites which would put up time series data. Likely early test client software includes SpecView, IRAF, DIS, VOPlot, and many others via the FITS interface.

Simple Image Access (SIA)

SIA has been in use now for over a year and it is time for a new version. Changes are needed to update the interface to reflect developments in the underlying technology over the past year, and to add new functionality to address issues found in the first round of applications.

The technology upgrades affect mainly the query response table (QRT). The changes under consideration include:

- Upgrade the query interface to add UCD1+ UCDs for all values in the response VOTable.
- Add UTYPE tags for all SIA interface elements.
- Revise the image metadata to use the emerging data model standards as for SSA. This includes uniform data characterization, coverage, dataset identification, and so forth.

Functionality enhancements include:

- Logical name proposal (R.Williams). This would provide a way to say that a subset of the images in the QRT are in some sense the “same” image, e.g., differing only in the file format or the spectral band.
- Ranking proposal (A.Rots). This would add a means to grade candidate images in the query response with the “best match” datasets receiving the highest grade or ranking.

F.Bonarrrel and P.Osuna have suggested that a mechanism be added to allow data providers to extend the SIA VOTable query response with non-standard metadata. This is already possible in a limited way by adding nonstandard FIELD extensions to the base table. The proposed new mechanism would add more extensive metadata in additional RESOURCE elements following the main QRT. This could take the form of either additional table elements, or structured data of some form. It was agreed that a subgroup would define some use cases and propose an extension mechanism. This is a good opportunity to experiment with mechanisms for extending VOTable.

Catalog Access

In Boston we saw major interest in adding general query language capabilities to the DAL services. The first step will be to add SkyNode as a new DAL service to replace the older cone search. More generally, we need to integrate query language capabilities (ADQL) into all the DAL services. This would take the form of extending the query interface to provide both a simple parameter (forms) based interface as at present, plus a more general ADQL interface to provide syntactic queries as well as support for advanced VO functionality such as regions. Only the query part of a DAL service would be affected. The query response would be the same for both parametric and syntactic queries, as would data model based data access. Virtual data generation can be supported by virtualization of the “table” being queried.

For example, integrating ADQL into SSA would mean adding a second ADQL-based query method. The simple parameter-based query would continue to be available, and the query response would be the same in both cases. Data access would be unaffected, returning a SED, spectrum, or time series conformant to the SSA data model, in any of the defined representations.

Our original DAL service architecture from the Cambridge meeting in 2003 defined both a set of type-specific interfaces (for catalogs, images, spectra, etc.) as well as a general “dataset” query. ADQL is exactly what is needed to implement the general dataset query. This service would provide a mechanism to search for any type of data. The query

response would return only general dataset metadata, and would include an indication of the type of data available.

Roadmap

Near term priorities:

- Complete SSA working draft. Initial service and client implementations. This is required in July 2004 to support the NVO summer school and for the 2005 IVOA demos.
- Second version of SIA. Prototype a DAL query extension mechanism to allow data provider-specific metadata to be returned.
- Define architecture and roadmap for integrating ADQL into DAL services (by late 2004).

By 2005 we should have stable basic versions of the SkyNode, SIA, and SSA services supporting all the major classes of astronomical data. More work on these will be needed of course, particularly in the areas of data model integration and web services support, but new priorities for 2005 should include specification of a next generation of all the DAL services integrating ADQL capabilities, plus a start on reference code for DAL including a data access framework and reference code implementing the client and server side of each DAL service. I don't think we can deploy something of the complexity of ADQL in DAL without a service framework and reference code.