

International VIRTUAL OBSERVATORY

# Data Access Layer Working Group Pune Workshop Summary

Doug Tody  
National Radio Astronomy Observatory

# SSA Data Model

- **Data model**
  - basic SSA data model is in good shape
  - dimensional equation looks promising for flux unit conversions
  - requires follow-up study for possible inclusion in SSA data model
- **Data model specification**
  - specification still incomplete, needs more detail
  - enumerate and rank allowable values, e.g., units and types
  - specify field names for all data model elements
  - specify what is required and what is optional

# SSA Data Representation

- **FITS**
  - format is binary table with (possibly variable length) array fields
  - no FITS ASCII tables
  - if text output is desired simple text format is preferred
- **VOTable**
  - format is SED header plus one table per segment
  - basic UTYPE usage determined but scoping rules need further study
  - SSA will use UCD1+
  - specify extension mechanism for data provider-supplied metadata
- **XML**
  - usage for basic metadata is very clean
  - mechanism for passing data points needs further study
- **Text**
  - will be used only for spectra and time series
  - CSV probably not adequate as cannot fully specify column metadata
  - format will be simple keyword header followed by tabular data

# SSA Query response table

- **Format**

- VOTable with elements of the SSA data model mapped to field groups
- Format similar to that used to represent a dataset in VOTable
- major difference is segment is a row not a table

- **Unresolved Issues**

- different metadata for SEDs than for spectrum or time series?
- do we include segment information for SEDs, or only SED descriptor?



# SIA Topics

- **Image Metadata**
  - metadata in main table will be greatly improved in SIA V1.1
  - will use new component data models as for SSA
- **Metadata extension proposal**
  - provides means to extend main table with additional metadata
  - basic proposal accepted without objection
  - impressive that a prototype test has already been conducted
  - further refinement will continue in subgroup
- **Asynchronous data staging**
  - needed for SIA, but a more general solution is required
  - DAL will work with GWS to determine how to do this
  - opportunity to integrate a number of grid technologies
  - asynchronous services, messaging, authentication, VOStore
  - SIA interface must hide underlying technology
  - must not compromise basic SIA usage or implementation

# ADQL Integration into DAL

- **Interface**
  - unified query mechanism with virtual tables looks very promising
  - prototyping required to learn how to use SQL with virtual data
- **Architecture**
  - in DAL terms ADQL will represent an alternative query mechanism
  - query response remains the same except for projection
  - data access, staging, etc., is unaffected
  - use of data models in DAL is unaffected
- **Protocol**
  - URL will be used for simple param = value queries
  - POST or SOAP will be used for passing encoded ADQL
  - whether or not DAL service interface should be based on SkyNode is TBD

# DAL Roadmap (Pune)

- **SSA**
  - V0.9 (target November 2004)
    - priority is to support initial test implementations, rather than interface completeness
    - develop support for time series, virtual data (e.g. SEDs)
  - V1.0 (target May 2005)
    - priority is interface completeness and reliability
    - for user service and client implementations
- **SIA**
  - V1.1 (target May 2005)
    - DM/Metadata updated (characterization etc.) as for SSA
    - limited feature support (logical name, ranking)
    - metadata extension mechanism
  - V2.0 (timescale TBD)
    - add support for spectral and time series data cubes
    - extended data model, data model-based data representation
    - including XML as well as FITS
- **ADQL integration into DAL**
  - focus on SkyNode and prototype effort between now and May 2005
  - require stable, reliable SIA, SSA before integrating ADQL
- **Reference implementations and framework**
  - in design and prototype phase currently