IVOA Data Access Layer
SSA Protocol Developments

Doug Tody (NRAO/NVO/IVOA)

SSA/DAL Interface

- Operations
  - queryData
    - Parameter-based interface most fundamental for DAL
    - Query language interface (ADQL) can be added later
      - Especially important for things like generic dataset queries
  - getData
    - URL-based currently
    - Possible: getData(AccessDID, [protocol-options])
      - e.g., data format, compression; persistent virtual data
  - stageData
    - Asynchronous data staging, VOSpace transports, bulk data
  - getCapabilities
    - Client queries, resource metadata

- Distributed Computing Platform
  - HTTP GET/POST currently
  - SOAP later; document-based still important

SSA/DAL Interface

- **Basic service elements**
  - Multiple operations
    - Share same service endpoint (base-URL)
    - REQUEST parameter defines operation

  - Protocol versioning
    - Service can support multiple interface versions
    - may share same service endpoint (baseURL)
    - Explicit runtime version control necessary for robustness
    - registry is a cache

  - VERSION parameter

- **HTTP GET operations**
  - \<baseURL\>?VERSION=1.0&REQUEST=queryData&PARAM=Value&...

SSA/DAL Interface

- **Basic service elements**
  - Parameters
    - Operation undefined if multiple instances
    - Parameter names are case-insensitive
    - Parameter values are case-sensitive unless otherwise specified
    - Numeric types as per XML Schema; sexagesimal not permitted

- **List-Structured Parameters**
  - [lower]'/'[upper]['/'step] ',' ... [';' <qualifier>] ...
  - Qualifiers apply to entire list
    - e.g., BAND="1E-6/3E-7;source"

Concepts

- **Data Derivation**
  - Where did this data come from?

  - Data Source
    - Original source of the data
      - survey, pointed, theory, artificial, composite

  - Creation Type
    - How the dataset returned by the service was created
      - archival, cutout, filtered, mosaic, projection, spectral extraction, catalog extraction

Concepts

- **Provenance**
  - Needed to describe virtual data returned by services

  - DataID.CreatorDID
    - dataset identifier assigned by creator

  - Curation.PublisherDID
    - dataset identifier assigned by publisher

  - Provenance = Dataset identifier + CreationType
    - e.g., CreationType="spectral extraction", CreatorDID=<cube>

Concepts

- **Association**
  - Logically associate records (datasets) in the query response
  - For SSA these are spectra
  - For complex data (generic dataset) can associate any datasets

- **Data Model**
  - Association.Type   Type of association (PARAM, ID-REF)
  - Association.ID   All records in assoc share same ID
  - Association.Key   Unique key tells how they differ

- **Extension metadata**
  - Use ID-REF to link to extension metadata for further info

- **Examples**
  - Echelle, multiband images, multiple data formats, data cubes with derived data products, spectra of cluster of galaxies, HST association, etc.

Input Parameters

- **POS**
  - Generalize to support coord frames other than ICRS
  - Needed to support galactic coords, solar and planetary data

  - Use list-structured format
    - POS='x,y[,z][;<frame>]
      - e.g., POS=12.0,22.3;GALACTIC_I
    - ICRS remains the default

- **SIZE**
  - unchanged

Input Parameters

- ## BAND
  - Generalize to permit both numeric and named bandpasses
    - e.g., BAND="1E-7/3E-6" or BAND="J"
  - Named bandpasses are determined by prior query to service
  - DataID.Bandpass attribute used to define named bandpass

- ## TIME
  - Unchanged (range list, ISO 8601 format)

- ## FORMAT
  - Only minor changes
  - Propose using MIME-type parameterization for format subtypes
  - Not normally needed as we have FORMAT=native

Optional Input Parameters

#	Parameter	Sample value	Unit	Datatype
1	APERTURE	0.00028 (=1″)	DDEG	double
2	**SPECRES**	2000	$d\lambda/\lambda$	double
3	**SPATRES**	0.05	DDEG	double
4	TIMERES	86760 (=1yr)	seconds	double
5	SNR	5.0	dimensionless	double
6	REDSHIFT	1.3/3.0	dimensionless	double
8	TARGETNAME	mars		string
7	TARGETCLASS	star		string

Optional Input Parameters

#	Parameter	Sample value	Unit	Datatype
8	**PUBDID**	ADS/col#R5983		string
9	**CREATORDID**	ivo://auth/col#R1234		string
10	**COLLECTION**	SDSS-DR4		string
11	**TOP**	20	dimensionless	int
12	**TOKEN**	50		string
13	**MTIME**	2005-01-01/2006-01-01	ISO 8601	string
14	**COMPRESS**	true		boolean
15	**RUNID**	<string>		string

Query Response

- **Query Metadata**
  - Query.Score   Scoring algorithm
  - Query.Token   Page through large queries

- **Association Metadata**
  - Association.Type, ID, Key

- **Access Metadata**
  - Access.Reference, Format, Size

- **Dataset/Spectrum Metadata**
  - SSA query response and Spectrum DM have been unified
    - Dataset, DataID, Curation, Target, Derived, Char, CoordSys

Query Response

- **General Dataset Metadata**

– Dataset.Type	Spectrum, TimeSeries, SED, etc.
– Dataset.DataModel	Data model (Spectrum V1.0)
– Dataset.Length	Number of data points in Spectrum
– Dataset.TimeAxisSI	SI factor and dimensions
– Dataset.SpectralAxisSI	SI factor and dimensions
– Dataset.FluxAxisSI	SI factor and dimensions

Query Response

- **Dataset Identification Metadata**

– DataID.Title	One-line description of dataset (String)
– DataID.Collection	Collection name (shortName)
– DataID.Creator	Creator of dataset (String)
– DataID.CreatorID	Identifier for VO Creator (URI)
– DataID.CreatorDID	Dataset ID assigned by creator (URI)
– DataID.CreatorLogo	URL for Creator logo
– DataID.Contributor	Contributor (may be multiple instances)
– DataID.Date	Date dataset created or last modified
– DataID.Version	Version of dataset instance (String)
– DataID.Instrument	Instrument which produced the data
– DataID.Bandpass	Spectral bandpass, e.g., a filter name
– DataID.DataSource	Original source of data
– DataID.CreationType	How was dataset created

Query Response

- **Curation Metadata**

– Curation.Publisher	Publisher (String)
– Curation.PublisherID	Identifier for VO Publisher (URI)
– Curation.PublisherDID	Dataset ID assigned by a publisher
– Curation.Rights	Restrictions on data
– Curation.Date	Date dataset created or last modified
– Curation.Version	Published version (String)
– Curation.Reference	URL or Bibcode for documentation
– Curation.ContactName	Contact name (String)
– Curation.ContactEmail	Contact email (Strin

Some SSA Implementations

- **Reference implementations**
  - VOClient
  - Client-side of DAL services (also registry etc.)
  - dalclient, eventually also dalserver Java libraries
  - Spectrum DM
  - Implements Spectrum DM as a Java class

- **Services**
  - JHU spectral service
  - MAST spectra

- **Applications**
  - VOSpec (display+), TSAP (theory data)
  - ESO/VOTech (SINFONI IFU data)
  - STScI/Specview (Java analysis client)
  - IRAF/splot,etc. (General data analysis)

- **Verification services**
  - NCSA

- **Others?**