

INTERNATIONAL VIRTUAL OBSERVATORY ALLIANCE
US National Virtual Observatory

IVOA Data Access Layer SIAP V2.0 Key Topics and Generic Dataset

D. Tody (NVO, NRAO)
A. Richards (Jodrell), F. Bonnarel (CDS)

SIAP V2.0 in Strasbourg

- **Topics**

- Status, recent activities
- Scope revisited
- Key SIAPV2 Topics
- Cube Use Cases (from Anita)
- Query Interface
- Access Functionality
- Grid Capabilities
- Roadmap

Recent Activities

- **Focus on selected key topics**
 - Mainly discussion among AR, FB, JS, DT in recent months
 - Anita Richards has joined as our radio data cube expert
- **Community involvement/input**
 - Mostly concerns radio data cube handling at this point
 - Getting organized for review and implementations
 - Need community involvement where cubes are involved
 - European AIDA initiative study group (Anita et.al.)
 - In US: primarily ALMA, EVLA, Arecibo

Scope Revisited

- **Primary capabilities**

- 2D images, cubes, grid
 - cube support required for radio community
 - grid required for some image generation cases, scalability

- **Strategy**

- Interface and model is N-D
- An individual service capability may however be 2D
- Basic interface and function comparable to SIAV1
 - but looks more like SSA, follows DAL2 profile and model

Key Topics

- **Status**
 - This has been our main focus for past few months
- **Study Areas**
 - Polarization, WCS/Mapping (incl FITS time),
 - Image Characterisation
 - Access functionalities (mainly cubes; reduction etc)
 - Service interface
- **Status**
 - Most topics understood well enough to proceed with V1
 - Largest remaining issues concern cube access functionalities

SIAP extension use cases

AIDA comments (AMSR)

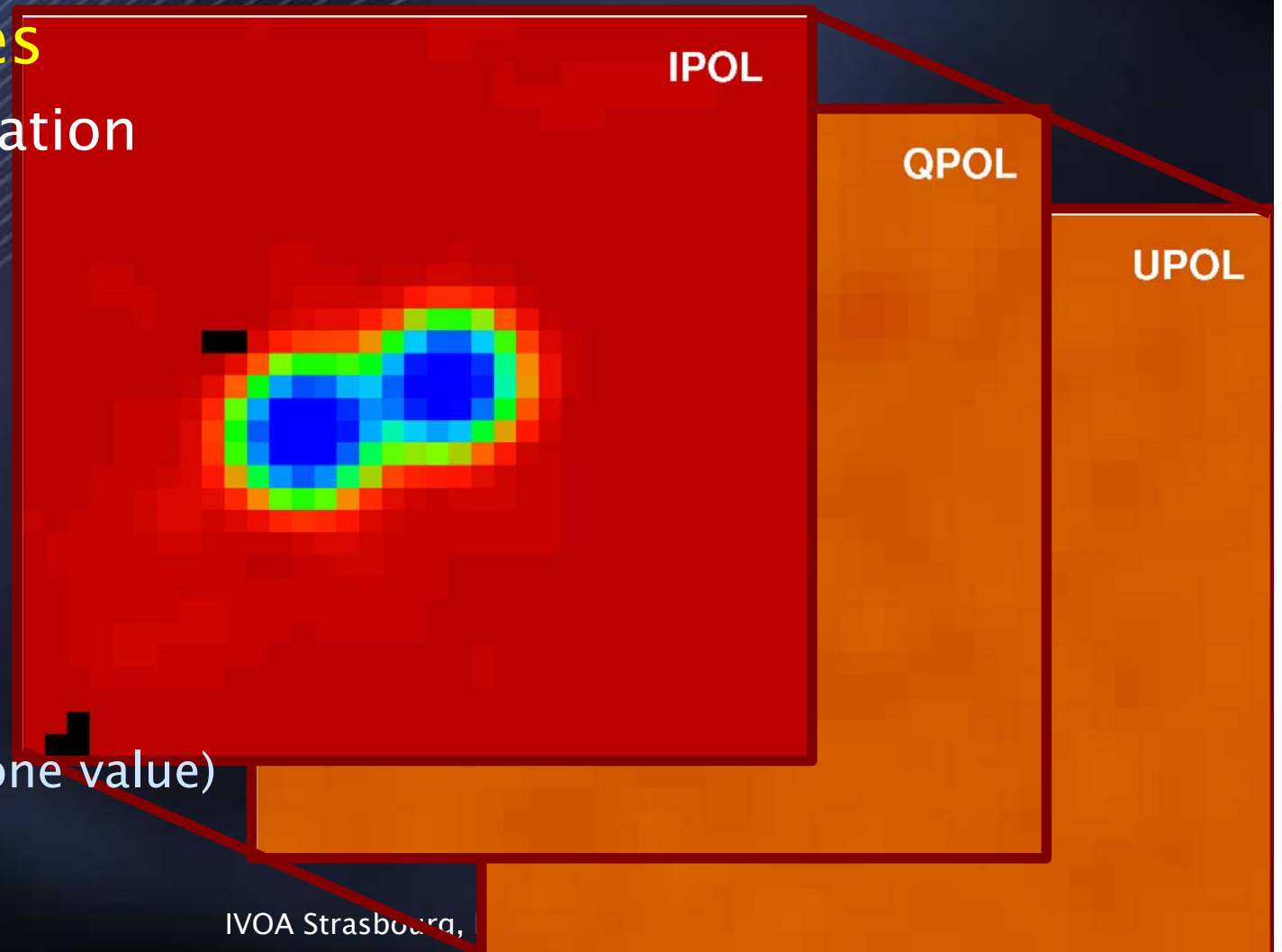
- **NVSS cubes**

- 3 polarization planes

- I
 - Q
 - U

- 4 Axes

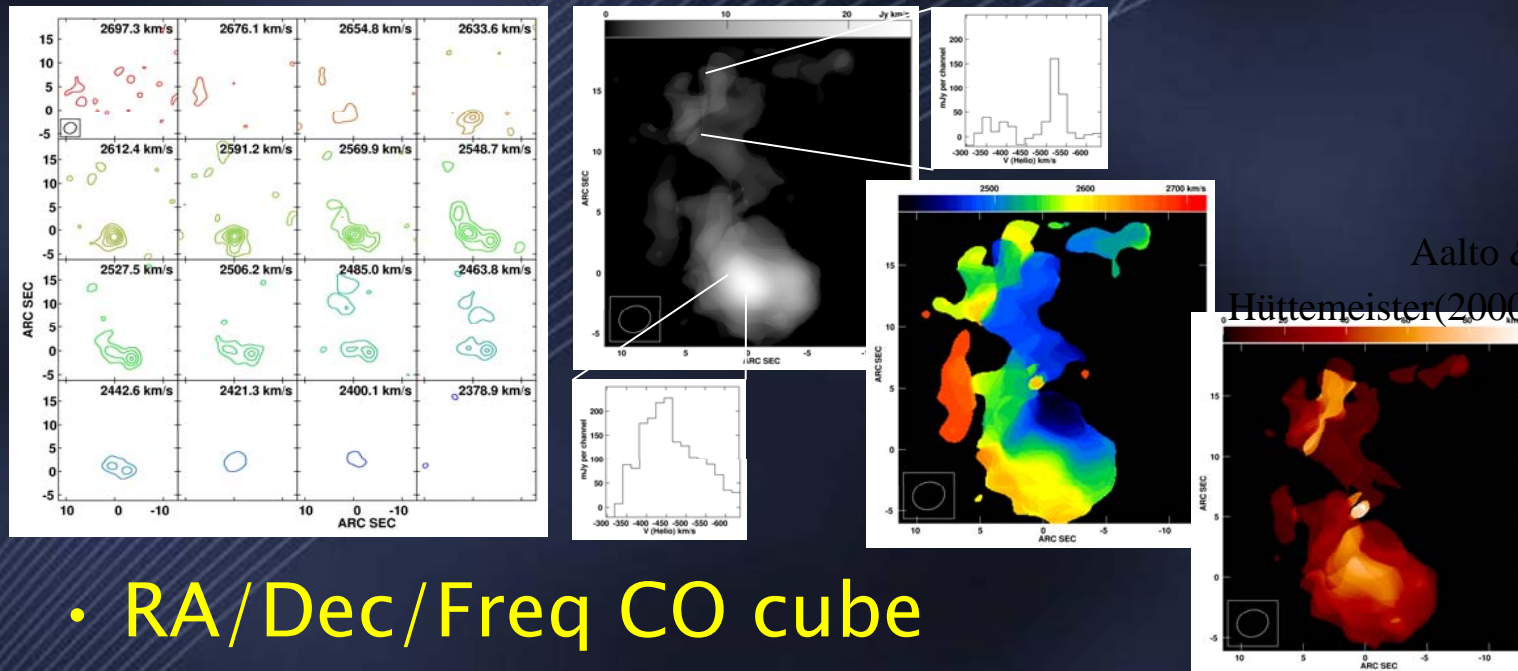
- RA
 - Dec
 - Stokes
 - Freq. (one value)



Spectral cubes

- 2 Spatial axes (RA, Dec., Glat, Glon etc.)
- Same observable (e.g. Jy/beam)
 - SNR etc. varies
- Spectral axis $N \times \delta$
 - Frequency, wavelength, energy
 - If $\delta\nu$ is linearly spaced, $\delta\lambda$ isn't
 - Velocity – need convention & line rest ν or λ or E
 - May be many lines in the cube
- Possible additional axes
 - Time, polarization...

ALMA/IRAM use case



- RA/Dec/Freq CO cube
 - Convert to velocity (LSR, radio convention)
 - Cutouts, simple squashes – VO tools?
 - Smoothed spectra, moments with noise cut-off
 - Specialised server-side pipeline controlled via UWS

Query Interface

- **POS,SIZE**

- Used for both searching and for access
 - people often think this is only for searching/discovery
- Choice of rectangle is mainly for access
 - defines "ideal image" footprint on the sky
 - used for simple cutouts and image generation
- Multi-position query needed here as with TAP

- **REGION**

- New parameter proposed for all DAL2 interfaces (STC-S)
- Used only for searching or spatial constraint
- Provides a general search region capability

Query Interface

- **BAND**

- Searching Semantics

- defines spectral coverage of data we are looking for
 - fixed wavelength units, or band name as in resource metadata
 - may also need to search by velocity interval?

- Access Semantics

- only an issue for spectral data cubes
 - defines cutout or filter on spectral axis
 - probably need to use native units of cube

- **VELOCITY, REDSHIFT**

- Proposed for conversions of spectral axis
 - Probably out of scope for V1; other approaches possible as well

Query Interface

- **TIME**

- Searching Semantics
 - ISO or MJD time or time range, fixed units
- Access Semantics
 - only an issue for time cubes
 - defines cutout or filter on time axis
 - units probably fixed at MJD in native time system of cube

- **POL**

- Searching Semantics
 - any data, only polarized data would be most typical
- Access Semantics
 - select (cutout) polarizations to return

Query Interface

- **SPECRES or SPECRP**

- Defines minimal spectral resolution or resolving power
- Only used for searching
- For SSA we chose SPECRP
- May not be natural choice for radio spectral data cubes

- **SPATRES**

- Defines minimal spatial resolution
- Refers to observed signal, not pixel size

- **TIMERES**

- Defines minimal time resolution (unit? – days or seconds)

Query Interface

- **TARGETNAME**
 - As for SSA. Required for non-positional data.
- **FLUXLIMIT**
 - Defines minimum allowable image sensitivity
 - SSA uses SNR, but this does not work in general for images
 - Need a measure of limiting flux such as Jy/unit-area
 - Anita suggests Jy/beam as close to what we need, but this is too radio-specific
- **Calibrations**
 - ASTCALIB, WAVECALIB, TIMECALIB, FLUXCALIB
 - Standard, as in SSA.

Query Interface

- **Curation**
 - Standard (dataset identifiers etc.)
- **Access Params**
 - Standard (MAXREC, MTIME, FORMAT, etc.).
- **Cube/Image Generation Parameters**
 - Specify geometry and WCS of output image.
 - One of our key topics.

Access Functionality

- **Status**
 - Another of our key topics.
 - Basics are understood, but some functionality such as reduction (squashing an axis) are still not scoped.
- **Possible Functionality**
 - Whole image, cutout, resample, reduce, transform
- **Cutout**
 - Range of POS, SIZE, BAND, TIME, POL

Access Functionality

- **Resample (various options)**
 - Specify WCS and geometry of desired output image
- **Reduce**
 - An axis gets smaller or goes away (reduced to single value)
 - Issue is what algorithms to support
 - sum, mean, spectral index, fractional polarization, POLA, etc.
 - potentially many options here...
 - May also incorporate filtering of spectral or time axis
- **Transform**
 - Probably out of scope for first version, but still under discussion
 - Examples are convert freq to velocity, Stokes to POLI, POLA, etc.

Grid Capabilities

- **Status**

- Should be easy after this is done in TAP

- **Async/UWS**

- Highest priority; required for advanced use cases
- Based upon concept of stageData operation
 - POST of params to standard UWS /async endpoint
- Job produces virtual images as described by QueryData
 - probably use access reference as ID tag
 - Any number of images can be produced in one job
- A standard acref GET can be used to retrieve each final image
 - Putting images in a co-located VOspace also possible

Roadmap

- **Status**

- Still working toward first working draft
- Hard to schedule more precisely until we have this
- Splinter meeting in Strasbourg

- **Tasks**

- Define strawman interface
 - discuss and agree within author team
 - review and iterate within DAL WG
- Produce V0.1 WD
 - discuss and iterate
- Initial Implementations
 - reference, prototype

Roadmap

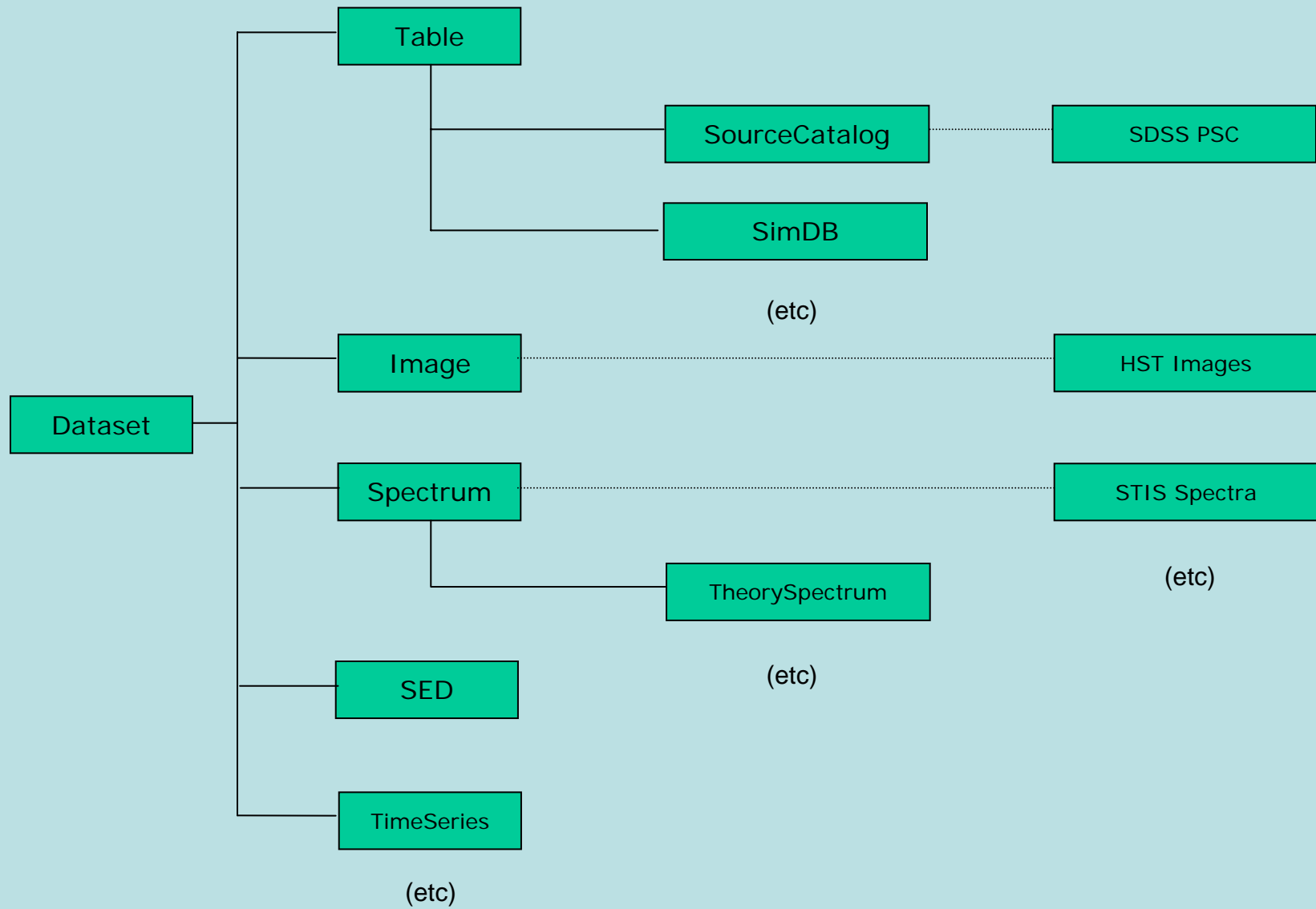
- **Initial Implementations**

- Reference implementations of 2D service
 - Should be quite straightforward
 - Basically what we have now updated to DAL2

- **Prototype of cube access capabilities**

- Both service and client (viewer) implementations require
 - e.g., Aladin and CASA viewer (ALMA) on client side
 - use largely existing software for cube access on server

Generic Dataset (GDS)



Generic Dataset

Top Level Class

Specialized Class

Instrumental Data

<i>Parameter</i>	<i>Sample value</i>	<i>Physical unit</i>	<i>Datatype</i>
POS	52,-27.8	degrees; defaults to ICRS	string
SIZE	0.05	degrees	double
BAND	2.7E-7/0.13	meters	string
TIME	1998-05-21/1999	ISO 8601 UTC	string
FORMAT	votable	-	string

Parameter	Sample value	Unit	Req	Datatype
SPECRP	2000	$\lambda/d\lambda$	REC	double
SPATRES	0.05	degrees	REC	double
TIMERES	31536000 (=1yr)	seconds	OPT	double
TARGETNAME	mars		OPT	string
TARGETCLASS	star		OPT	string
ASTCALIB	absolute		OPT	string
WAVECALIB	absolute		OPT	string
FLUXCALIB	relative		OPT	string
PUBDID	ADS/col#R5983		REC	string
CREATORID	ivo://auth/col#R1234		REC	string
COLLECTION	SDSS-DR5		REC	string

GDS Schema

- **Datasets**

- GDS data model (a type of index table)
- Usually a subset with local additions
- Records may be associated using Association model

- **Links**

- Table of data links
- Uses ID from Dataset table as a key

Associations

- Usage

- Normally these are stored as table fields
 - In VOTable, constant values can be PARAMs
- A record can belong to more than one association

- Model

- Association.Type
 - Type of association (a string)
- Association.ID
 - Unique ID identifying the association instance
- Association.Key
 - Unique key different for each element of association

Data Links

- **Usage**

- Link an object (e.g. table record) to some other object
- An object can have a number of such links
- Some standard link types are defined but set is extensible
- GDS uses a table representation but others are possible (RDF)

- **Model (Strawman)**

- LinkId Used as the link reference in Dataset table
- LinkType Type of link (standard or local extension)
- URL URL of data object, service, etc.

- **Sample Link Types**

- File of some kind (table, image, spectrum, etc., or custom)
- A standard service which can be used to access the data
- A custom service which can be used to access the data
- HTML page, etc.

GDS Queries

- **Strategy**
 - GDS would normally be stored in tables using schema
 - Hence it can be queried with TAP
 - TAP param query (PQL) extension to support GDS query
 - ADQL could also be used once we have UTYPE support in TAP
 - this is essentially a TAP subclass
- **Custom Index Tables**
 - Site may have arbitrary local index tables
 - Can extend these with GDS metadata or data links
 - Use standard VO tools to browse and access data
- **Usage**
 - Client tools can browse and query data, follow links,
 - invoke services and display the result, download data,
 - link data to a vospace, invoke pipeline, etc.