

TOPCAT TAP Client Improvements

Mark Taylor (Bristol)

IVOA Interop
Sesto

17 June 2015

`$Id: toptap.tex,v 1.20 2015/06/17 08:45:14 mbt Exp $`

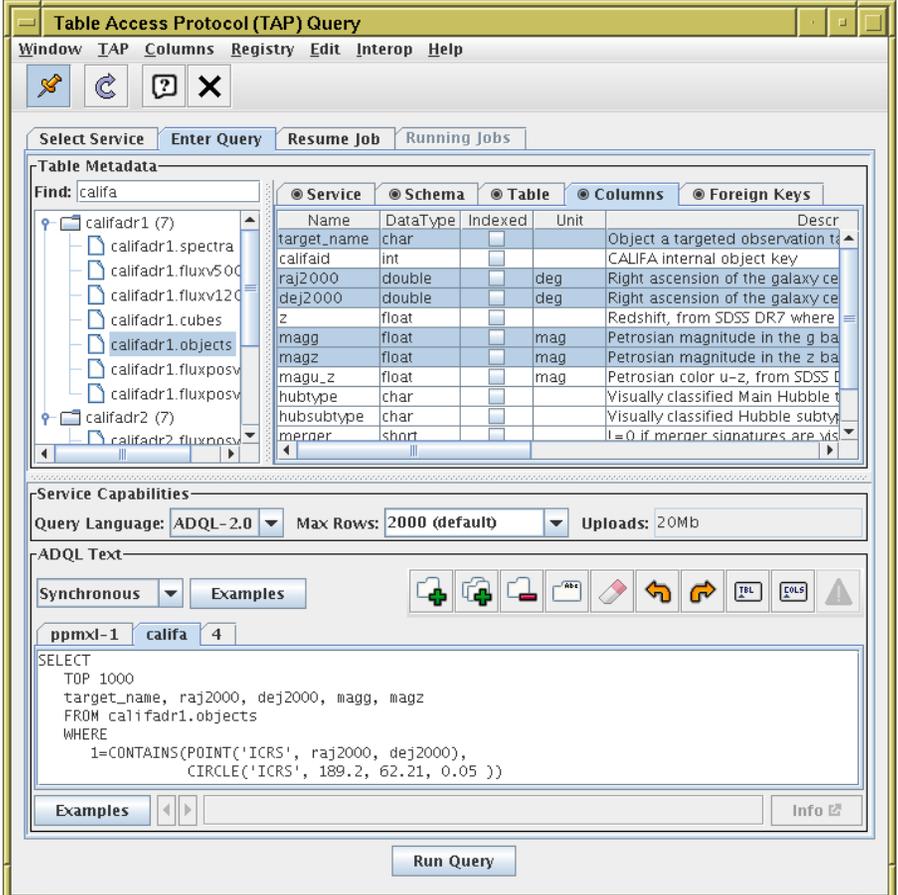
Outline

- Summary of improvements
- Demo
- Implementation feedback
- Status
- Outstanding issues

Improvements

New in this (pre-)release:

- Improved metadata acquisition: pluggable and scalable to many tables
- Improved metadata display GUI: scalable to many tables
- More per-service and per-table metadata displayed
- More examples, including service-provided ones
- Better ADQL editing panel
- Query result FORMAT negotiation
- Quick-look query mode



The screenshot displays the Table Access Protocol (TAP) Query interface. The main window is titled "Table Access Protocol (TAP) Query" and contains several panels:

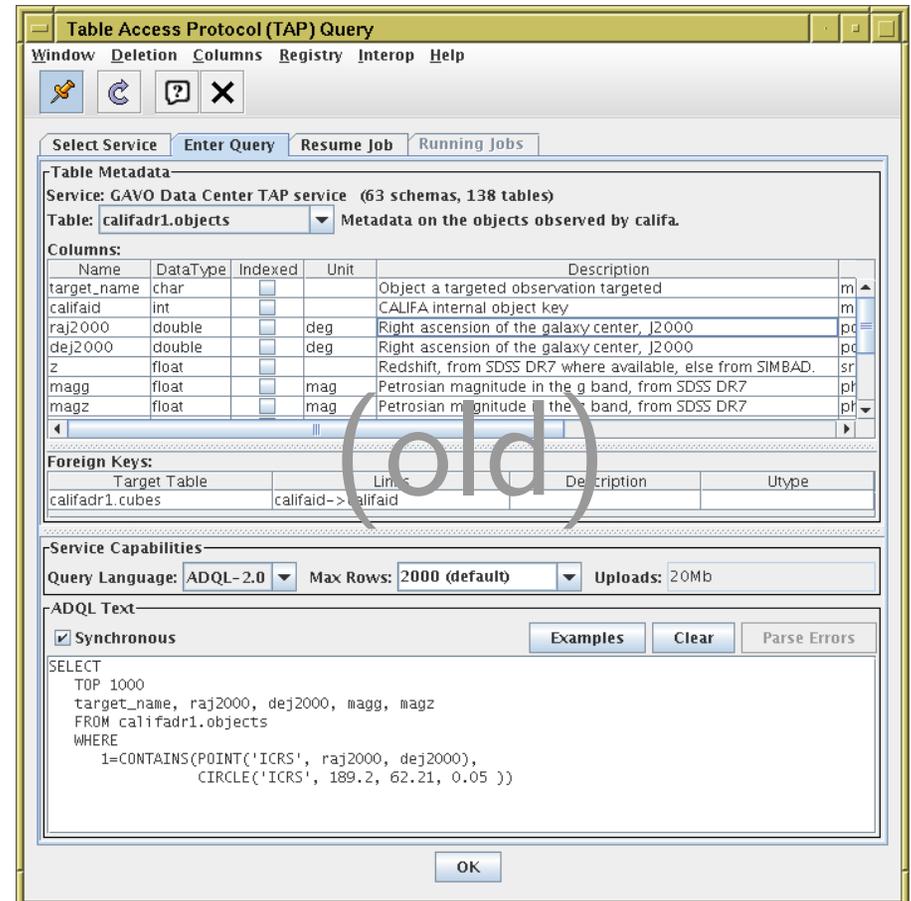
- Table Metadata:** A tree view on the left shows a hierarchy of tables under the "califa" service. The right pane shows a table with columns: Name, DataType, Indexed, Unit, and Description. The table lists various astronomical parameters like target_name, califa_id, ra_j2000, etc.
- Service Capabilities:** A section below the metadata showing "Query Language: ADQL-2.0", "Max Rows: 2000 (default)", and "Uploads: 20Mb".
- ADQL Text:** A text editor area containing an ADQL query. The query is:

```
SELECT TOP 1000 target_name, raj2000, dej2000, magg, magz FROM cali1fadr1.objects WHERE 1=CONTAINS(POINT('ICRS', raj2000, dej2000), CIRCLE('ICRS', 189.2, 62.21, 0.05 ))
```
- Buttons:** "Run Query" at the bottom, and "Examples" and "Info" buttons on the left and right of the ADQL text area.

Improvements

New in this (pre-)release:

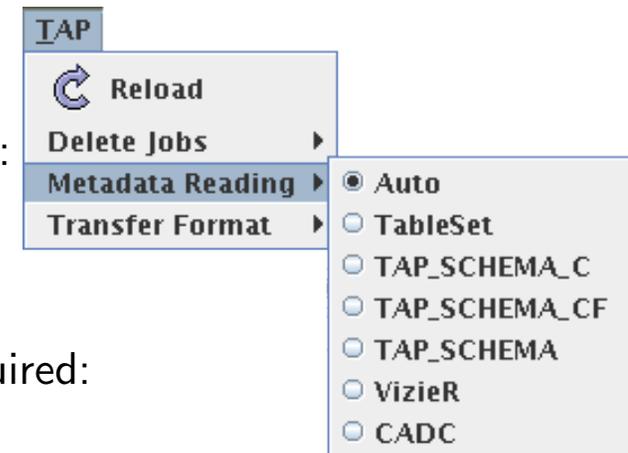
- Improved metadata acquisition: pluggable and scalable to many tables
- Improved metadata display GUI: scalable to many tables
- More per-service and per-table metadata displayed
- More examples, including service-provided ones
- Better ADQL editing panel
- Query result FORMAT negotiation
- Quick-look query mode



Metadata acquisition

Pluggable metadata acquisition options:

- 1-stage: All metadata (schemas, tables, columns, keys) up front:
 - ▷ from standard `/tables` endpoint
 - ▷ from `TAP_SCHEMA` queries
- 2-stage: Tables and schemas up front, columns and keys as required:
 - ▷ from `TAP_SCHEMA` queries
 - ▷ from CADC-style multi-stage `/tables` endpoint (proposed VOSI 1.1)
 - ▷ from VizieR-style multi-stage `/tables` endpoint (non-standard)
- Auto
 - ▷ Chooses one of the above (e.g. by using total column count)



2-stage options are scalable

- Workable with VizieR-scale service (30k tables, 400k columns, $\lesssim 10$ seconds initial load)
- Some care required when scheduling/threading column metadata requests

Additional tweaks made in some cases

- Attempt to quote not-delimited-but-should-be column/table identifiers

Metadata display

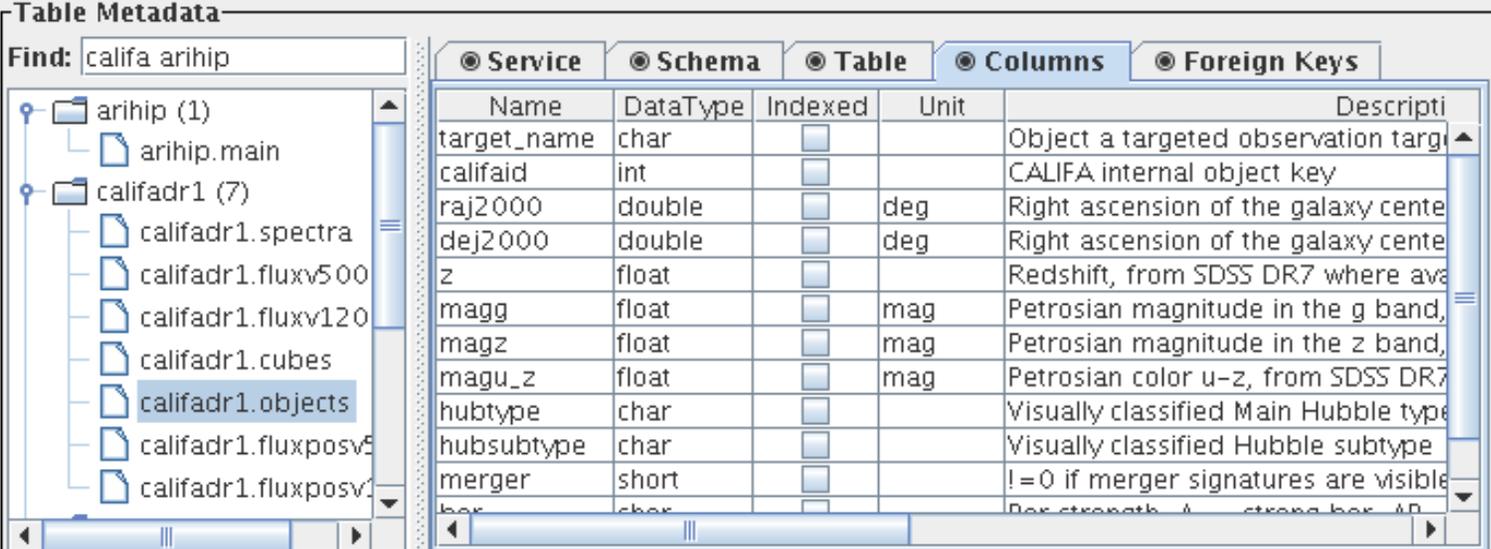
- Better scalability

- JTree not JComboBox
- Tables displayed hierarchically by schema
- Instant search by table name

→ *usable with TAPVizieR!*

- More information displayed

- Service metadata from registry (description, reference URL, ...)
- TAPRegExt language features (supported DMs, UDFs, geometry functions, ...)
- Full table and schema metadata (descriptions)



The screenshot shows a software interface titled "Table Metadata". On the left, there is a "Find:" search box containing "califa arihip". Below it is a hierarchical tree view showing a folder "arihip (1)" containing "arihip.main", and another folder "califadr1 (7)" containing several sub-tables, with "califadr1.objects" selected. On the right, there are tabs for "Service", "Schema", "Table", "Columns", and "Foreign Keys". The "Columns" tab is active, displaying a table with the following data:

Name	Data Type	Indexed	Unit	Description
target_name	char	<input type="checkbox"/>		Object a targeted observation target
califaid	int	<input type="checkbox"/>		CALIFA internal object key
raj2000	double	<input type="checkbox"/>	deg	Right ascension of the galaxy center
dej2000	double	<input type="checkbox"/>	deg	Right ascension of the galaxy center
z	float	<input type="checkbox"/>		Redshift, from SDSS DR7 where available
magg	float	<input type="checkbox"/>	mag	Petrosian magnitude in the g band, from SDSS DR7
magz	float	<input type="checkbox"/>	mag	Petrosian magnitude in the z band, from SDSS DR7
magu_z	float	<input type="checkbox"/>	mag	Petrosian color u-z, from SDSS DR7
hubtype	char	<input type="checkbox"/>		Visually classified Main Hubble type
hubsubtype	char	<input type="checkbox"/>		Visually classified Hubble subtype
merger	short	<input type="checkbox"/>		!=0 if merger signatures are visible
box	char	<input type="checkbox"/>		Box strength A strong box AD

Metadata display

- Better scalability

- JTree not JComboBox
- Tables displayed hierarchically by schema
- Instant search by table name

→ *usable with TAPVizieR!*

- More information displayed

- Service metadata from registry (description, reference URL, ...)
- TAPRegExt language features (supported DMs, UDFs, geometry functions, ...)
- Full table and schema metadata (descriptions)

Table Metadata
Service: GAVO Data Center TAP service (63 schemas, 138 tables)
Table: Metadata on the objects observed by califa.

Columns:

Name	Data Type	Indexed	Unit	Description	
target_name	char	<input type="checkbox"/>		Object a targeted observation targeted	meta. ▲
califaid	int	<input type="checkbox"/>		CALIFA internal object key	meta. ≡
raj2000	double	<input type="checkbox"/>	deg	Right ascension of the galaxy center, J2000	pos.e
dej2000	double	<input type="checkbox"/>	deg	Right ascension of the galaxy center, J2000	pos.e
z	float	<input type="checkbox"/>		Redshift, from SDSS DR7 where available, else from SIMBAD.	src.re ▼

Foreign Keys:

Target Table	Links	Description	Utype
califadr1.cubes	califaid->califaid		

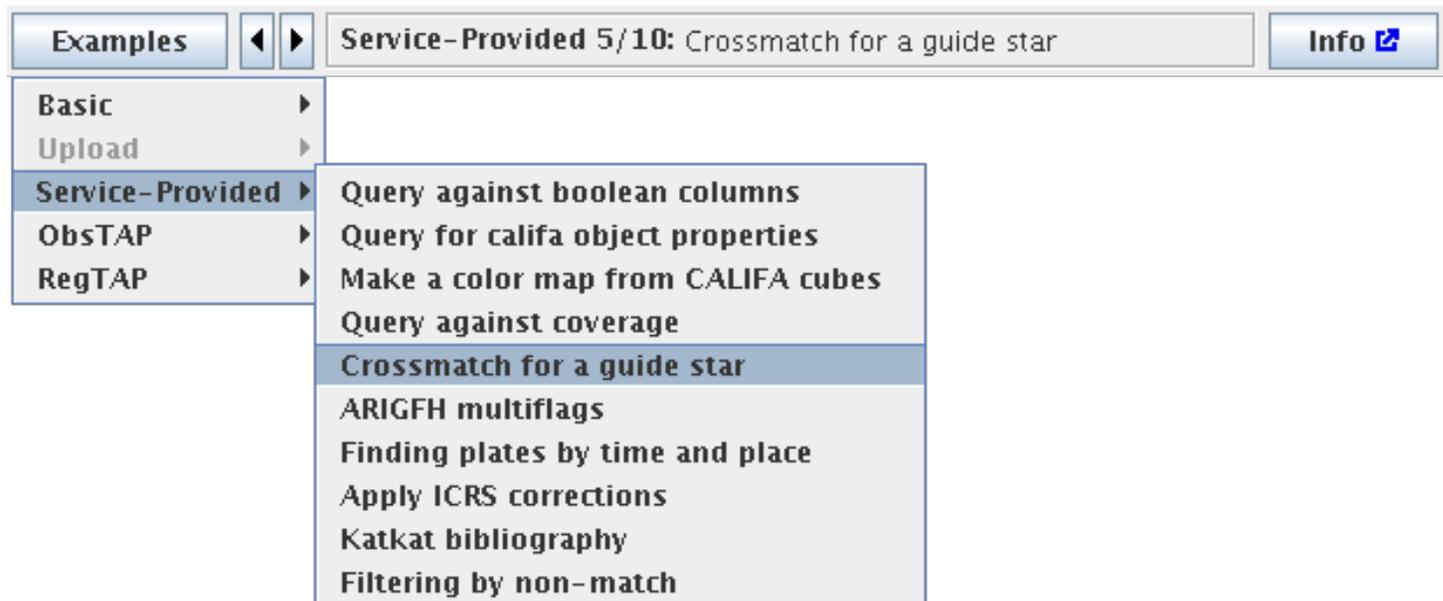
Examples

More ADQL example queries:

- Service-specific examples from service DALI/TAP [/examples](#) endpoint (XHTML+RDFa)
- Canned RegTAP and ObsTAP examples if data models declared

Example browsing:

- Buttons to cycle through example lists
- Link to example documentation URLs (opened in browser)



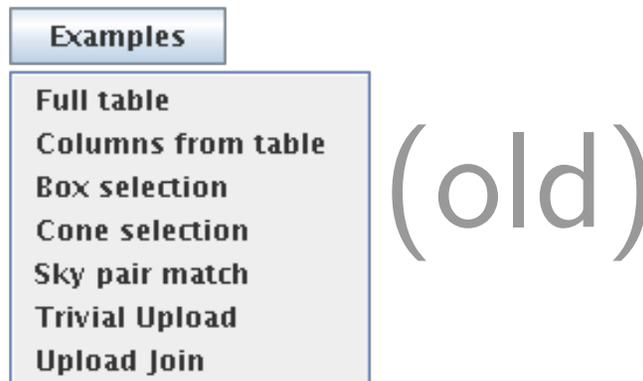
Examples

More ADQL example queries:

- Service-specific examples from service DALI/TAP [/examples](#) endpoint (XHTML+RDFa)
- Canned RegTAP and ObsTAP examples if data models declared

Example browsing:

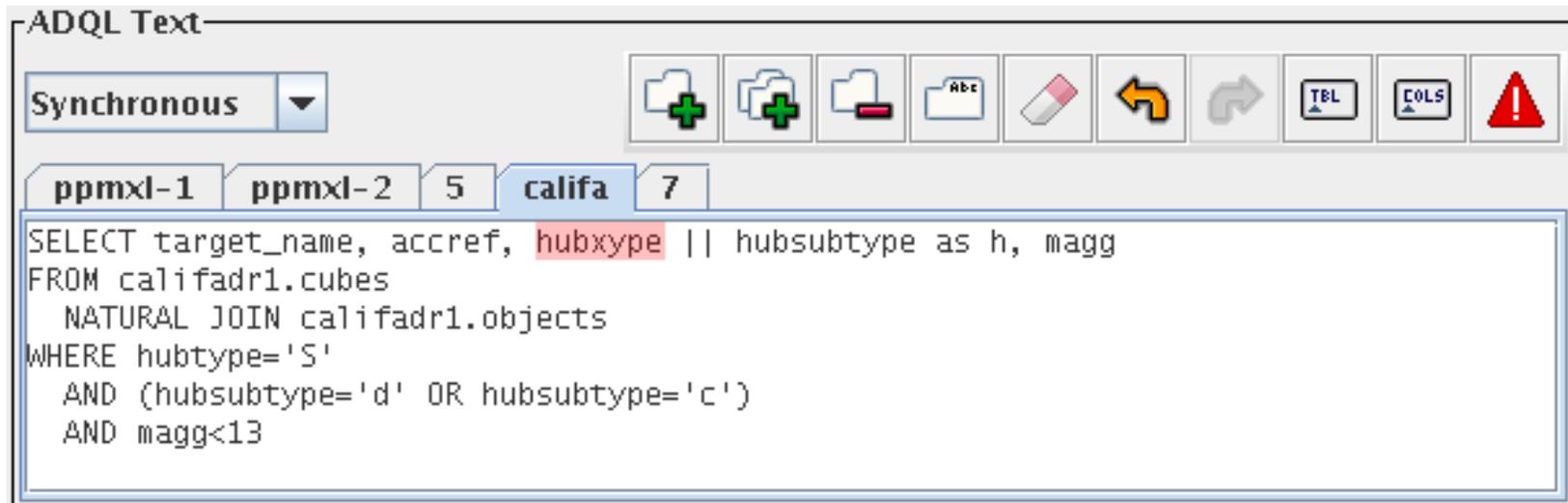
- Buttons to cycle through example lists
- Link to example documentation URLs (opened in browser)



ADQL Editing

More functions added to ADQL editing window

- Multiple tabs allow multiple queries on the go
- Text edit Undo/Redo actions (buttons or Ctrl-[Shift]-Z)
- Convenience buttons for inserting selected table/column names
- Version 1.3 of CDS/ARI ADQL lib for metadata-aware ADQL parsing
UDF-aware parsing coming soon?



ADQL Editing

More functions added to ADQL editing window

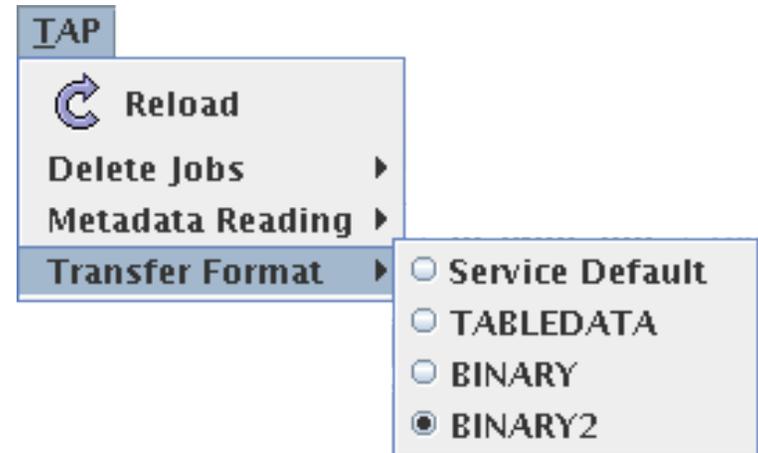
- Multiple tabs allow multiple queries on the go
- Text edit Undo/Redo actions (buttons or Ctrl-[Shift]-Z)
- Convenience buttons for inserting selected table/column names
- Version 1.3 of CDS/ARI ADQL lib for metadata-aware ADQL parsing
UDF-aware parsing coming soon?

```
-ADQL Text
 Synchronous
Examples Clear Parse Errors
SELECT target_name, accref, hubxype || hubssubtype as h, magg
FROM califadr1.cubes
  NATURAL JOIN califadr1.objects
WHERE hubtype='S'
  AND (hubssubtype='d' OR hubssubtype='c')
  AND magg<13
```

Result Format Negotiation

Uses FORMAT parameter for result transfer format selection

- Default to BINARY2 if service provides it
- User may choose preferred format by menu
- Performance may be affected, results not much



Use HTTP Content-Coding for HTTP compression

- Can deliver big reductions in bandwidth use
- See my talk in GWS I

Demo

The screenshot displays the Table Access Protocol (TAP) Query interface. The window title is "Table Access Protocol (TAP) Query". The menu bar includes "Window", "TAP", "Columns", "Registry", "Edit", "Interop", and "Help". The toolbar contains icons for a pin, refresh, help, and close.

The interface is divided into several sections:

- Select Service**: A tabbed interface with "Enter Query" selected.
- Table Metadata**: A search box contains "califa". A tree view on the left shows a folder "califadr1 (7)" with sub-items like "califadr1.objects" selected. A table on the right shows metadata for the selected table.
- Service Capabilities**: A section with "Query Language: ADQL-2.0", "Max Rows: 2000 (default)", and "Uploads: 20Mb".
- ADQL Text**: A section with a "Synchronous" dropdown, an "Examples" button, and a text area containing an ADQL query. Below the text area are "Examples" and "Info" buttons.

The "Table Metadata" table is as follows:

Service	Schema	Table	Columns	Foreign Keys		
		Name	Data Type	Indexed	Unit	Descr
		target_name	char	<input type="checkbox"/>		Object a targeted observation t
		califa_id	int	<input type="checkbox"/>		CALIFA internal object key
		raj2000	double	<input type="checkbox"/>	deg	Right ascension of the galaxy ce
		dej2000	double	<input type="checkbox"/>	deg	Right ascension of the galaxy ce
		z	float	<input type="checkbox"/>		Redshift, from SDSS DR7 where
		magg	float	<input type="checkbox"/>	mag	Petrosian magnitude in the g ba
		magz	float	<input type="checkbox"/>	mag	Petrosian magnitude in the z ba
		magu_z	float	<input type="checkbox"/>	mag	Petrosian color u-z, from SDSS D
		hubtype	char	<input type="checkbox"/>		Visually classified Main Hubble t
		hubsubtype	char	<input type="checkbox"/>		Visually classified Hubble subtyp
		merger	short	<input type="checkbox"/>		! = 0 if merger signatures are vis

The ADQL Text area contains the following query:

```
SELECT
TOP 1000
target_name, raj2000, dej2000, magg, magz
FROM califadr1.objects
WHERE
  1=CONTAINS(POINT('ICRS', raj2000, dej2000),
    CIRCLE('ICRS', 189.2, 62.21, 0.05 ))
```

Implementation Feedback

General comments:

- TAP and related standards are complex, but mostly provide what's required
- RegTAP is a huge improvement on RI 1.0 (I can now rely on registry queries)
- TAP service quality is patchy (implementation errors, missing functionality)

Status

Available in pre-release

- Working
- User documentation TBD
- Available for standalone/third-party use (TOPCAT application classes not required)
- TAP service client fairly complete
- TAP service discovery still poor
- Your feedback is welcome, especially service providers
- Public release (TOPCAT v4.2-4?) in next few weeks?

ftp://andromeda.star.bris.ac.uk/pub/star/topcat-full_tap.jar

Outstanding Issues

TBD:

- **Better discovery UI for (useful) TAP services**
- Some remaining ADQL library issues
 - ▷ Periods in table names
 - ▷ UDF validation
- Standards under discussion
 - ▷ DALI/TAP [/examples](#) endpoint
 - ▷ Scalable [/tables](#) endpoint (VizieR vs. CADC)
- UWS 1.1 (*no services*)
- Use HTTP Content-Coding? (see my talk from GWS I)
- Fixup machinery for non-identifier names
- Multiple TAP window instances per TOPCAT?

Open Questions

- Authentication?
- Anything else I should be doing?
- Complaints/concerns from service providers?