

DataLink Implementation Feedback

Mark Taylor (Bristol)

IVOA Interop
Victoria

31 May 2018

`$Id: dlfeedback.tex,v 1.9 2018/05/31 16:49:48 mbt Exp $`

Outline

Implementation feedback relating to DataLink

- DataLink document
- Service Descriptor context
- Service Descriptor content
- Service Descriptor positioning
- Row Correspondance
- DataLink usage patterns

Context:

- New DataLink-related functionality in TOPCAT
- See talk in Apps II

The screenshot shows the TOPCAT(1): DataLink View window. It contains a table with the following data:

ID	access_url	service_def
6	ivo://org.gavo.dc/~?califa...	procsvc0
7	ivo://org.gavo.dc/~?califa...	procsvc1
8	ivo://org.gavo.dc/~?califa...	procsvc2
9	ivo://org.gavo.dc/~?califa... http://dc.zah.uni-heidelb...	
10	ivo://org.gavo.dc/~?califa... http://dc.zah.uni-heidelb...	
11	ivo://org.gavo.dc/~?califa... http://dc.zah.uni-heidelb...	
12	ivo://org.gavo.dc/~?califa... http://dc.zah.uni-heidelb...	
13	not-a-real-identifier	NotFc

Below the table, the 'Row Link Type' is 'Fixed Access URL'. The 'Row Detail' section shows the following information:

- access_url:** http://dc.zah.uni-heidelberg.de/getproduct/califa/datadr3/COMB/
- content_type:** image/jpeg
- content_length:**
- description:** A preview for the dataset.
- semantics:** #preview
- URL:** http://dc.zah.uni-heidelberg.de/getproduct/califa/datadr3/COMB/ARP220.CC

The 'URL' section also includes a 'Type' dropdown set to 'IMAGE', an 'Action' dropdown set to 'View image internally', and an 'Invoke' button.

- DataLink document

- Had to read it lots of times
- Talks about distinct things:
 - ▷ {Links} service
 - ▷ {Links}-response format VOTable
 - ▷ Service Descriptor VOTable RESOURCE elements
- One way of putting these together is discussed in the document
- But they can be used in other ways (e.g. standalone SDs), seem to be less carefully thought through

Service Descriptor Context

- Service Descriptors come in two different contexts:
 - *Either* referenced by {links}-response document
or standalone in any old VOTable
 - These two contexts work differently:
 - ▷ Standalone SDs apply implicitly to every table row
 - ▷ {links}-document SDs apply only to rows that reference them explicitly
 - This complicates VOTable processing
 - ▷ need to check specifically if document looks like a {links} response to work out how to treat SDs
 - See [Datalink service descriptor usage](#) DAL thread

Service Descriptor Content

- Service Descriptors don't have much standard metadata
 - Only `accessURL`, `standardID`, `resourceIdentifier` are discussed in REC-DataLink-1.0
 - Doesn't really matter for {links}response services, but for standalone SDs, there's not much for clients or users to go on
 - ▷ ... especially if there is more than one

```
<RESOURCE type="meta" utype="ad hoc:service">
```

```
  <PARAM name="accessURL" value="http://example.com/cone/query" .../>
```

```
  <PARAM name="standardID" value="ivo://ivoa.net/std/cone" .../>
```

```
  <PARAM name="resourceIdentifier" value="ivo://example.com/cone" .../>
```

```
  <GROUP name="inputParams">
```

```
    ...
```

```
  </GROUP>
```

```
</RESOURCE>
```

Service Descriptor Content

- Service Descriptors don't have much standard metadata
 - Only `accessURL`, `standardID`, `resourceIdentifier` are discussed in REC-DataLink-1.0
 - Doesn't really matter for {links}response services, but for standalone SDs, there's not much for clients or users to go on
 - ▷ ... especially if there is more than one

```
<RESOURCE type="meta" utype="adhoc:service" name="Cone Search">
  <DESCRIPTION>Positional search</DESCRIPTION>
  <PARAM name="accessURL" value="http://example.com/cone/query" .../>
  <PARAM name="standardID" value="ivo://ivoa.net/std/cone" .../>
  <PARAM name="resourceIdentifier" value="ivo://example.com/cone" .../>
  <PARAM name="contentType" value="application/x-votable+xml" .../>
  <GROUP name="inputParams">
    ...
  </GROUP>
</RESOURCE>
```

- Suggested additions:
 - ▷ Use standard VOTable `name` attribute and `DESCRIPTION` element in DataLink SD examples
 - ▷ Define a new `contentType`-named PARAM

Service Descriptor Positioning

Table & associated Service Descriptor(s) are in sibling RESOURCES

- Table has `type="results"`, SD has `type="meta"`
- Sequence is unspecified
 - ▷ Difficult when streaming VOTables: may want to know SDs up front
- No explicit rule for multi-table VOTable documents (*maybe out of scope for DataLink?*)
 - ▷ Difficult when serializing multiple tables with SD metadata (TOPCAT save session)
 - ▷ ... this may be a bit out of the scope of the DataLink document
 - ▷ How to do it?

```
<VOTABLE>
  <RESOURCE type="results"><TABLE name="table1"> ... </TABLE></RESOURCE>
  <RESOURCE type="meta"> ... service descriptor ... </RESOURCE>
  <RESOURCE type="results"><TABLE name="table2"> ... </TABLE></RESOURCE>
</VOTABLE>
```

→ whose is the service descriptor?

Service Descriptor Positioning

Table & associated Service Descriptor(s) are in sibling RESOURCES

- Table has `type="results"`, SD has `type="meta"`
- Sequence is unspecified
 - ▷ Difficult when streaming VOTables: may want to know SDs up front
- No explicit rule for multi-table VOTable documents (*maybe out of scope for DataLink?*)
 - ▷ Difficult when serializing multiple tables with SD metadata (TOPCAT save session)
 - ▷ ... this may be a bit out of the scope of the DataLink document
 - ▷ How to do it?

```
<VOTABLE>
  <RESOURCE>
    <RESOURCE type="results"><TABLE name="table1"> ... </TABLE></RESOURCE>
    <RESOURCE type="meta"> ... service descriptor ... </RESOURCE>
  </RESOURCE>
  <RESOURCE>
    <RESOURCE type="results"><TABLE name="table2"> ... </TABLE></RESOURCE>
  </RESOURCE>
</VOTABLE>
```

→ Aladin doesn't find the tables 😞

Service Descriptor Positioning

Table & associated Service Descriptor(s) are in sibling RESOURCES

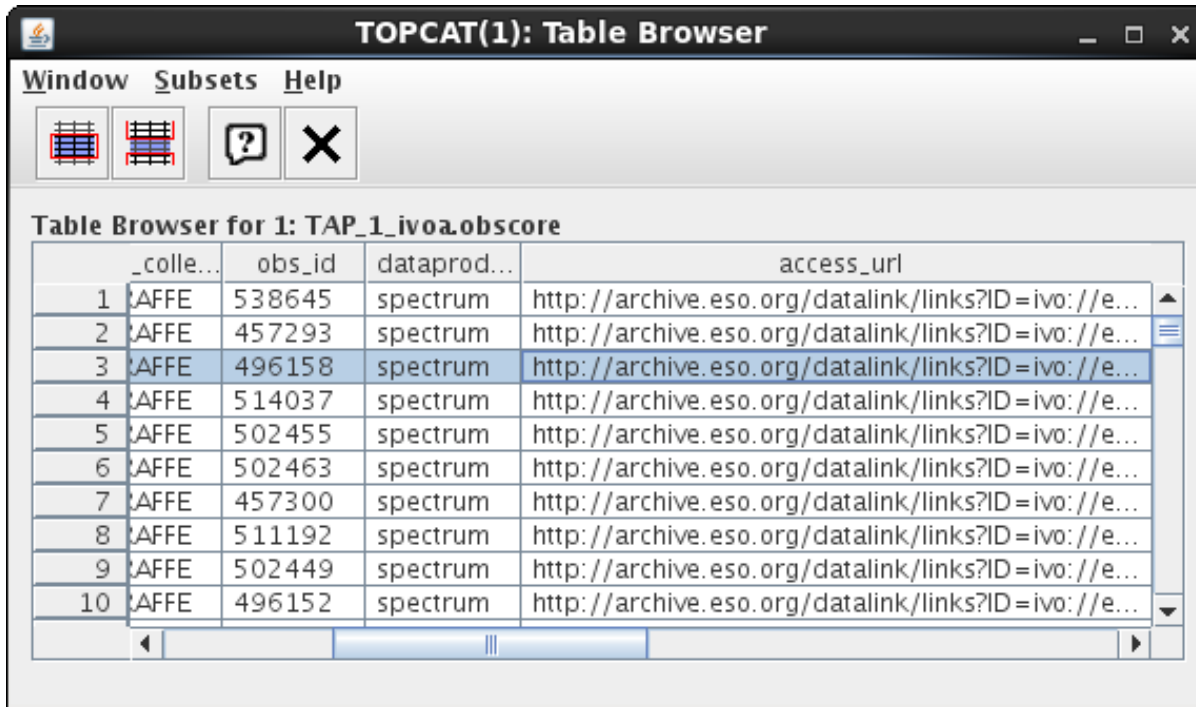
- Table has `type="results"`, SD has `type="meta"`
- Sequence is unspecified
 - ▷ Difficult when streaming VOTables: may want to know SDs up front
- No explicit rule for multi-table VOTable documents (*maybe out of scope for DataLink?*)
 - ▷ Difficult when serializing multiple tables with SD metadata (TOPCAT save session)
 - ▷ ... this may be a bit out of the scope of the DataLink document
 - ▷ How to do it?

```
<VOTABLE>
  <RESOURCE>
    <TABLE name="table1"> ... </TABLE>
    <RESOURCE type="meta"> ... service descriptor ... </RESOURCE>
  </RESOURCE>
  <RESOURCE>
    <TABLE name="table2"> ... </TABLE>
  </RESOURCE>
</VOTABLE>
```

→ Current solution in TOPCAT — doesn't break anything

Row Correspondance

- Would like to highlight “the same” row in related {links} tables



TOPCAT(1): Table Browser

Window Subsets Help

Table Browser for 1: TAP_1_ivoa.obscore

	_colle...	obs_id	dataprod...	access_url
1	AFFE	538645	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
2	AFFE	457293	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
3	AFFE	496158	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
4	AFFE	514037	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
5	AFFE	502455	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
6	AFFE	502463	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
7	AFFE	457300	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
8	AFFE	511192	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
9	AFFE	502449	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...
10	AFFE	496152	spectrum	http://archive.eso.org/datalink/links?ID=ivo://e...

Row Correspondance

- Would like to highlight “the same” row in related {links} tables

The image shows two overlapping windows from the TOPCAT software. The background window is 'TOPCAT(1): Table Browser' showing a table with columns for observation ID, data product, and URL. The foreground window is 'TOPCAT(1): Activation - View Datalink Table' showing a 'DataLink Table' with columns for semantics, description, and category. Row 6 is highlighted in both windows. Below the table in the foreground window, there is a 'Row Link Type' section with 'Fixed Access URL', a 'Row Detail' section with various metadata fields, and a 'URL' section with a dropdown menu set to 'FITS_IMAGE', an 'Action' dropdown set to 'View image internally', and an 'Invoke' button. The 'Result' field shows 'FAIL' with the message 'No responses'.

	_colle...	obs_id	dataprod...	
1	AFFE	538645	spectrum	http://archive.eso.org/...
2	AFFE	457293	spectrum	http://archive.eso.org/...
3	AFFE	496158	spectrum	http://archive.eso.org/...
4	AFFE	514037	spectrum	http://archive.eso.org/...
5	AFFE	502455	spectrum	http://archive.eso.org/...
6	AFFE	502463	spectrum	http://archive.eso.org/...
7	AFFE	457300	spectrum	http://archive.eso.org/...
8	AFFE	511192	spectrum	http://archive.eso.org/...
9	AFFE	502449	spectrum	http://archive.eso.org/...
10	AFFE	496152	spectrum	http://archive.eso.org/...

	semantics	description	eso_catego...
1	#this	Requested file	SCIENCE.SPECTRUM
2	#preview	PDF preview of the requested file	
3	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
4	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
5	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
6	#auxiliary	Ancillary file associated to the re...	ANCILLARY.PREVIEW
7	#progenitor	File used to generate the request...	
8	http://archive.eso.org/rdf/datali...	Release description of the collecti...	

Row Link Type
Fixed Access URL

Row Detail
access_url: http://archive.eso.org/datalink/links?ID=ivo://eso.org/ID?ADP.2015-04-13T10:11:05.923
content_type: application/fits
content_length: 69076
description: Ancillary file associated to the requested file
semantics: #auxiliary

URL: http://archive.eso.org/datalink/links?ID=ivo://eso.org/ID?ADP.2015-04-13T10:11:05.923

Type: **FITS_IMAGE** Action: **View image internally** **Invoke**

Result: FAIL No responses

Row Correspondance

- Would like to highlight “the same” row in related {links} tables

The image shows two windows from the TOPCAT software. The left window, titled 'TOPCAT(1): Table Browser', displays a table with columns for observation ID, data product, and URL. The right window, titled 'TOPCAT(1): Activation - View Datalink Table', shows a detailed view of a specific data link, including its semantics, description, and access URL. The 'View Datalink Table' window also includes a 'Row Link Type' section, a 'Row Detail' section with various metadata fields, and a 'URL' section with a dropdown menu for the link type and an 'Invoke' button.

Table Browser for 1: TAP_1_ivoa.obscure

	_colle...	obs_id	dataprod...	
1	AFFE	538645	spectrum	http://archive.eso.org/...
2	AFFE	457293	spectrum	http://archive.eso.org/...
3	AFFE	496158	spectrum	http://archive.eso.org/...
4	AFFE	514037	spectrum	http://archive.eso.org/...
5	AFFE	502455	spectrum	http://archive.eso.org/...
6	AFFE	502463	spectrum	http://archive.eso.org/...
7	AFFE	457300	spectrum	http://archive.eso.org/...
8	AFFE	511192	spectrum	http://archive.eso.org/...
9	AFFE	502449	spectrum	http://archive.eso.org/...
10	AFFE	496152	spectrum	http://archive.eso.org/...

DataLink Table

	semantics	description	eso_catego...
1	#this	Requested file	SCIENCE.SPECTRUM
2	#preview	PDF preview of the requested file	
3	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
4	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
5	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
6	#auxiliary	Ancillary file associated to the re...	ANCILLARY.PREVIEW
7	#progenitor	File used to generate the request...	
8	http://archive.eso.org/rdf/datali...	Release description of the collecti...	

Row Link Type
Fixed Access URL

Row Detail

access_url: http://archive.eso.org/datalink/links?ID=ivo://eso.org/ID?ADP.2015-04-13T10:12:41.540

content_type: application/fits

content_length: 51713

description: Ancillary file associated to the requested file

semantics: #auxiliary

URL: http://archive.eso.org/datalink/links?ID=ivo://eso.org/ID?ADP.2015-04-13T10:12:41.540

Type: **FITS_IMAGE** Action: **View image internally** **Invoke**

Result: FAIL No responses

Row Correspondance

- Would like to highlight “the same” row in related {links} tables

The image shows two windows from the TOPCAT software. The left window is the 'Table Browser for 1: TAP_1_ivoa.obscore', displaying a table with columns: _colle..., obs_id, dataprod..., and a URL column. The right window is 'TOPCAT(1): Activation - View Datalink Table', which shows a 'DataLink Table' with columns: semantics, description, and eso_catego. Below this table are sections for 'Row Link Type' (Fixed Access URL), 'Row Detail' (access_url, content_type, content_length, description, semantics), and a 'URL' section with a 'Type' dropdown (FITS_IMAGE), an 'Action' dropdown (View image internally), and an 'Invoke' button. The 'Result' is shown as 'FAIL No responses'.

	semantics	description	eso_catego
1	#this	Requested file	SCIENCE.SPECTRUM
2	#preview	PDF preview of the requested file	
3	#auxiliary	Ancillary file associated to the re...	ANCILLARY.MOSSPECT
4	#auxiliary	Ancillary file associated to the re...	ANCILLARY.PREVIEW
5	#progenitor	File used to generate the request...	
6	http://archive.eso.org/rdf/datali...	Release description of the collecti...	

	access_url	content_type	content_length	description	semantics
	http://archive.eso.org/datalink/links?ID=ivo://eso.org/ID?ADP.2015-04-13T10:	application/fits	50549	Ancillary file associated to the requested file	#auxiliary

URL: http://archive.eso.org/datalink/links?ID=ivo://eso.org/ID?ADP.2015-04-13T10:11:05.37C

Type: **FITS_IMAGE** Action: **View image internally** **Invoke**

Result: FAIL No responses

- How to identify which is the same?
- Row index? Look for similar (semantics, description) values?
- New column link_code(?), assessed for equality, might be nice

DataLink Usage

- DataLink URL column pattern:

(a.k.a. “DataLink Recognition Problem”)

- VOTable has a column with a URL referencing a DataLink table
- This looks like a “standard” /expected usage pattern
- ... but there’s no way to mark up the VOTable to indicate that’s what it is
- Hard for clients to present an intuitive/easy user interface

Hacky solution: `columnName.toLowerCase().contains("datalink")`

- see [Characterising URL columns](#) DAL thread

- Lots of freedom in usage

😊 Good for services

- ▷ can accommodate many different ways to distribute associated data

😞 Some problems for clients/users

- ▷ may not be obvious what usage pattern the data provider has in mind
- ▷ may be difficult to configure default/sensible/repeated use for a given “link”