Fig. 1

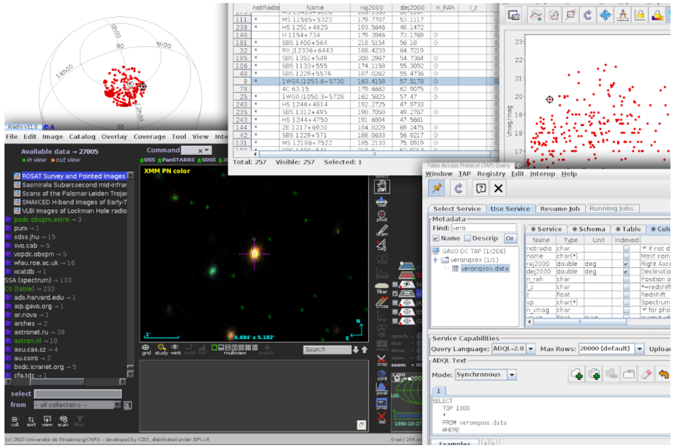
Fig. 2

1. The Virtual Observatory

Markus Demleitner
msdemlei@ari.uni-heidelberg.de

Since 2002, most astronomical data centres worldwide have collaborated in the IVOA[1] to make their holdings machine-consumable and uniformly findable.

Today, $\sim$ 25′000 data collections with many 10^8 datasets and several 10^{10} table rows are (almost) uniformly queriable through the Virtual Observatory (VO).

In this talk, I will give a brief idea of what we have built. Given the VO consists of many servers talking to many clients about a large variety of data, this cannot be more than a brief impression. Feel free to contact me with further questions.

(cf. Fig. 1)

2. First: A Demo

(cf. Fig. 2)

As a non-astronomer, it is probably difficult to reproduce this, but here is a short summary of what has been shown.

1 `https://www.ivoa.net`

1. Start TOPCAT[2]
2. Open the VO → TAP dialog, in there look for Quasar
3. Select the GAVO TAP service that has the veronqso schema, a catalogue of quasars, and hit "Use service"
4. Locate the veronqsos.data table and investigate its column metadata
5. Run a (trivial) query against that by letting TOPCAT come up with a sample query: Example → cone selection, then change the cone to use 189, 62 as the center and 15 degrees as the radius
6. Hit Graphics → Sky Plot
7. Hit Graphics → Plane Plot and plot the redshift z vs. the brightness Vmag
8. Click on points in the diagrams and watch TOPCAT update the selection in the other – that also works between VO-enbled applications.
9. Start Aladin[3] and in TOPCAT, do Views → Activation Actions and check "Send coordinates"
10. In Aladin, select DSS and zoom in to a field of view of a few arcminutes. Then in TOPCAT, click on an object with a Vmag of something like 20. In Aladin, you should see a very faint counterpart
11. In Aladin, navigate to Image → X-ray → XMM and add the XMM False Color plane (note how there are green and orange data collections: the green ones are the ones that may have data for the area shown)
12. Note how most quasars will be prominent in X-rays even if inconspicuous optically – these are strange beasts!

3. What did just happen?

What you saw was the co-operation of about a dozen software components – clients and services – that (conceptually) know nothing of each other. They cooperate because they all adhere to standards.

The rest of this talk highlights some of them relevant for the demo.

Find them all on http://ivoa.net/documents.

2 `http://www.star.bris.ac.uk/~mbt/topcat/`
3 `https://aladin.cds.unistra.fr/AladinDesktop/`

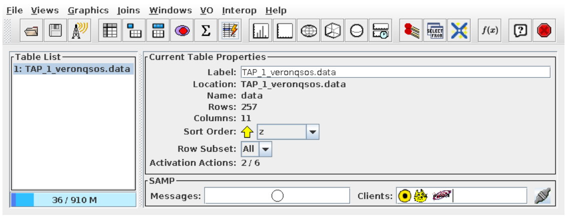

Fig. 3

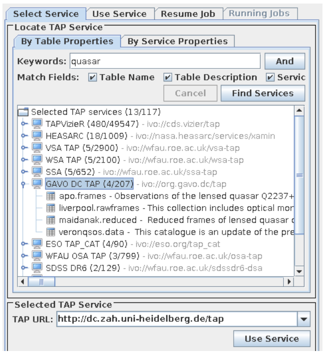

Fig. 4

4. TOPCAT

(cf. Fig. 3)

Table manipulation and operation of many VO protocols – evolved out of a pre-VO UK computational astronomy project (starlink). More[4].

There are other popular clients speaking VO protocols. A popular alternative is using pyvo within jupyter notebooks.

5. Discovery of TAP Services

(cf. Fig. 4)

I do discovery (here, by keyword) among all (∼70000) tables registered for queriability in the world, coming from ∼130 different services provided by ∼ 50 operators.

This works because of the VO Registry (VOResource[5], Registry Interfaces[6], VODataService[7], and others). Of course, all these standards enable a lot more than just a keyword query, and

4 http://www.star.bris.ac.uk/~mbt/topcat/

5 http://ivoa.net/Documents/VOResource/

6 http://ivoa.net/Documents/RegistryInterface

7 http://ivoa.net/documents/VODataService/

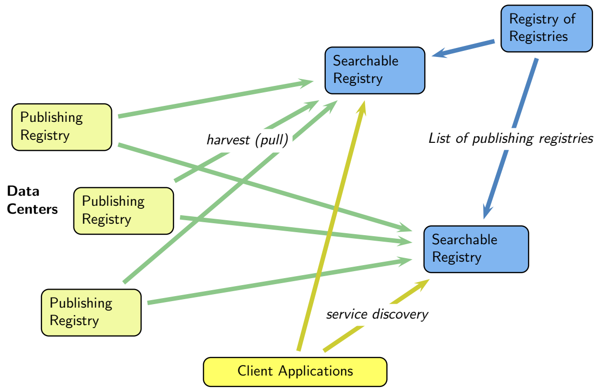

Fig. 5

behind the scenes, TOPCAT *does* a lot more than just a keyword query in order to come up with this interface.

6. VO Registry: Architecture

(cf. Fig. 5)

Only one minor central component (the Registry of Registries).

Each data provider runs a publishing registry, which is harvested by full registries.

Clients only talk to full registries (right now: GAVO, ESA, and NASA each run one).

7. VO Registry: Metadata Schema

(cf. Fig. 6)

Discovery is mostly through a relational mapping of the XML schema (RegTAP[8]) that also gives a high-level overview of the metadata schema.

Full Disclosure: the TOPCAT TAP discovery by default still uses a custom service called GloTS that re-publishes (mainly) registry metadata rather than RegTAP directly. This is one of those historical accidents that sometimes are hard to rectify in widely distributed systems.

8 http://ivoa.net/documents/RegTAP/

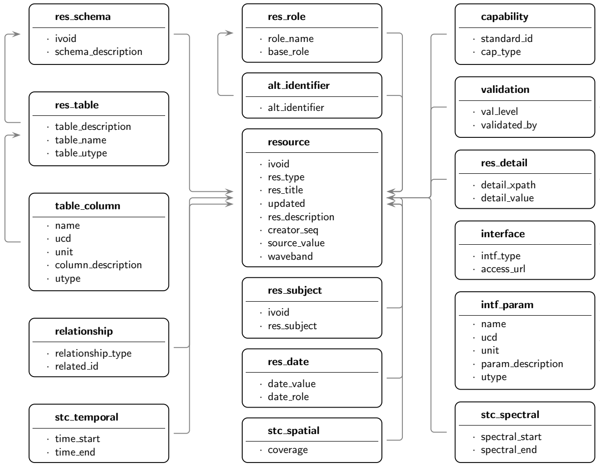

Fig. 6

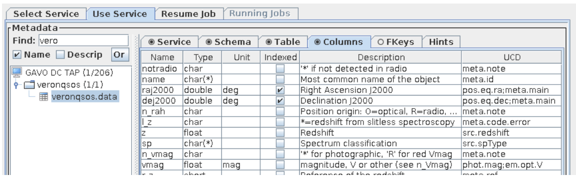

Fig. 7

8. TAP: Metadata

(cf. Fig. 7)

Without a special contract with the service, TOPCAT immediately shows table and column metadata including actionable units (VOUnit[9]) and formal physics (UCDs[10]).

That's because the Table Access Protocol TAP[11] offers (not one but two – ahem) methods to ask a service to provide them.

[9] http://ivoa.net/documents/VOUnits/

[10] http://ivoa.net/Documents/latest/UCD.html

[11] http://ivoa.net/Documents/TAP

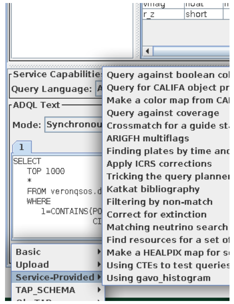

Fig. 8

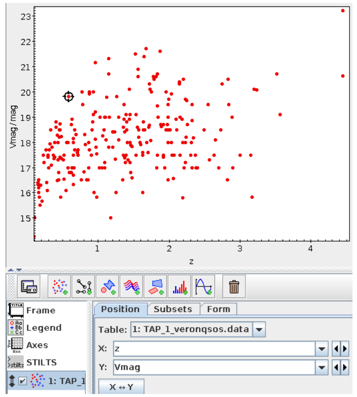

Fig. 9

9. TAP: Writing a Query

(cf. Fig. 8)

TAP services are queried in ADQL[12], a custom dialect of SQL that is *constant across all servers* regardless of the backend database (there are translators).

Astronomers do learn SQL, but they appreciate examples. TOPCAT produces some itself, others come from the service providers in XHTML+RDFa.

10. Metadata-Rich Results

(cf. Fig. 9)

Again, without any special contract, TOPCAT can produce meaningful, labelled plots from what it gets from the TAP service.

[12] http://ivoa.net/Documents/latest/ADQL.html


Fig. 10

That's thanks to VOTable[13], an XML-based table format designed to carry expressive metadata.

11. A VOTable

```
<VOTABLE xmlns="http://www.ivoa.net/xml/VOTable/v1.3" version="1.4">
  <DESCRIPTION>This catalogue contains 11358 (+2759) quasars (defined...
  </DESCRIPTION>
  <INFO name="legal" value="Data and table metadata come from..."/>
  <RESOURCE type="results">
    <INFO name="query" value="SELECT TOP 2 * FROM veronqsos.data"/>
    <INFO name="QUERY_STATUS" value="OK">Query successful</INFO>
    <INFO name="citation" value="1998yCat.7207....0V"/>
    <TABLE name="data">
      <FIELD ID="raj2000" datatype="double" name="raj2000"
        ucd="pos.eq.ra;meta.main" unit="deg">
        <DESCRIPTION>Right Ascension J2000</DESCRIPTION>
        <VALUES><MIN value="0.0117"/><MAX value="359.9971"/></VALUES>
      </FIELD>...
      <DATA><BINARY>
        <STREAM encoding="base64">IAAAAApT...=</STREAM>
</BINARY></DATA></TABLE></RESOURCE></VOTABLE>
```

This is of course abridged. Simple TAP can be run with just curl. To see the full table, run

```
curl -FLANG=ADQL -FQUERY="SELECT TOP 2 * FROM veronqsos.data" \
  http://dc.g-vo.org/tap/sync \
 xmlstarlet fo  less
```

12. Inter-Client Communication

(cf. Fig. 10)

To see how my result looks on the sky, I sent the table over from TOPCAT to Aladin[14], a client produced by CDS.

The two understood each other thanks to the Simple Application Messaging Protocol SAMP[15]. That's IPC with high-level, disciplinary primitives ("Send a VOTable", "Send the result of a Spectral query").

13 http://ivoa.net/Documents/VOTable

14 https://aladin.cds.unistra.fr/aladin.gml

15 http://ivoa.net/documents/SAMP/

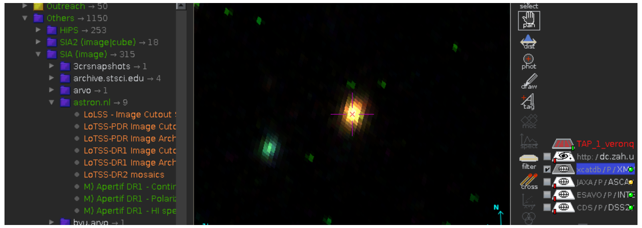
Fig. 11

13. Zooming through the Spectrum

(cf. Fig. 11)

Aladin has its own view on the Registry, where it marks resources that spatially cover the current image in green thanks to Multi-Order Coverage maps (MOCs[16]). Zooming and panning are enabled by a related standard called HiPS[17].

14. A Library

By the way, workflows developed in this way can usually be scripted, for instance using pyVO. Here's Python source code for loading the result from a TAP query into Aladin (using TAP, VOTable, SAMP, and more):

```
import pyvo

svc = pyvo.dal.TAPService("http://dc.g-vo.org/tap")
result = svc.run_sync("""
SELECT TOP 1000 *
   FROM veronqsos.data
   WHERE
      1=CONTAINS(POINT('ICRS', raj2000, dej2000),
                 CIRCLE('ICRS', -30, -20, 15 ))""")

with pyvo.samp.connection() as conn:
  pyvo.samp.send_table_to(conn,
    result.to_table(), "Aladin")
```

See PDF attachment(s): inpython.py

15. Parting Words

It took about 30 standards before the VO started to become halfway useful to astronomers. Crossing the death valley of writing standards that nobody would actually use was not easy.

Almost all of our standards somehow have astronomy coded in. But within each standard, usually only a small part (5% to 40%) actually is astronomy-specific. Can this be factored out?

If I had to start now: I'd compare OGC and VO standards and bootstrap from what they have in common.

16 http://ivoa.net/documents/MOC/

17 http://ivoa.net/documents/HiPS/20170519/