GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

PUBLIC DOMAIN

Bibliographic Interfaces in the VO

Markus Demleitner

IVOA Southern Spring Interop 2024, Malta Nov 14-17

Data Curation and Preservation IG

BibVO

Primary goal: "Populate ADS/SciX D-Links from VO resources."

Three scenarios:

- A VO resource x "is" data for publication y.
- A VO resource has dataset(s) $\vec{x}$ for publication y.
- A VO resource should be included into SciX's index as a publication.

Each scenario requires a different approach and has a BibVO section of its own.

VO Resource is Data for Article

That's been possible since day 1 of VOResource: Put a bibcode into *content/source*.

```
<content>
  <subject>astrostatistics-techniques</subject>
  <subject>milky-way-galaxy</subject>
  <subject>astronomical-simulations</subject>
  <description> This catalogue is a simulation
    ...</description>
  <source format="bibcode">2018PASP..130g4101R</source>
  <referenceURL>http://dc.g-vo.org/browse/gdr2mock/q</referenceURL>
```

content/source in ADS

VO Ressource has Datasets

This is the main part of BibVO: "Spectra $x_1, x_2, \ldots x_n$ that I serve are related to publication y".

ADS already had custom protocols with some providers of observatory bibliographies before. BibVO sect. 3 proposes a standard for that:

biblink-harvest

biblink-harvest

You run one or more endpoints without parameters which return records giving:

- bibref-ref: a paper id
- dataset-ref: a URI ADS should use for a D-link
- relationship: Cites (paper derives results from dataset) or IsSupplementedBy (paper produced dataset)
- bib-format (defaults to bibcode, otherwise as for content/source/@format)
- anchor-text (for link elements)
- cardinality (when you map to lists of datasets)

Serialisation and Discovery in biblink-harvest

Currently, this data is served in JSON:

```
[{'bib-ref ': '2020A&A...637A...4R',
'dataset-ref ': 'https://dc.g-vo.org/LP/org.gavo.dc/toss/q/line_tap',
  'relationship ': 'IsSupplementedBy'},
{'bib- ref ': '2012A&A...546A..55R',
'dataset-ref ': 'https://dc.g-vo.org/LP/org.gavo.dc/toss/q/line_tap',
  'relationship ': 'IsSupplementedBy'}]
```

Register endpoints delivering such data with capabilities with a standardId of

ivo://ivoa.net/std/bibvo#biblink-harvest-1.0

Note

There are discussions whether some details should change:

- https://github.com/ivoa/BibVO/issues/3 Should we use different relationships?
- https://github.com/ivoa/BibVO/issues/4 Should we be a bit more RDFish?

Don't let that hold you up: Your implementations will only have to change very little if we make changes here.

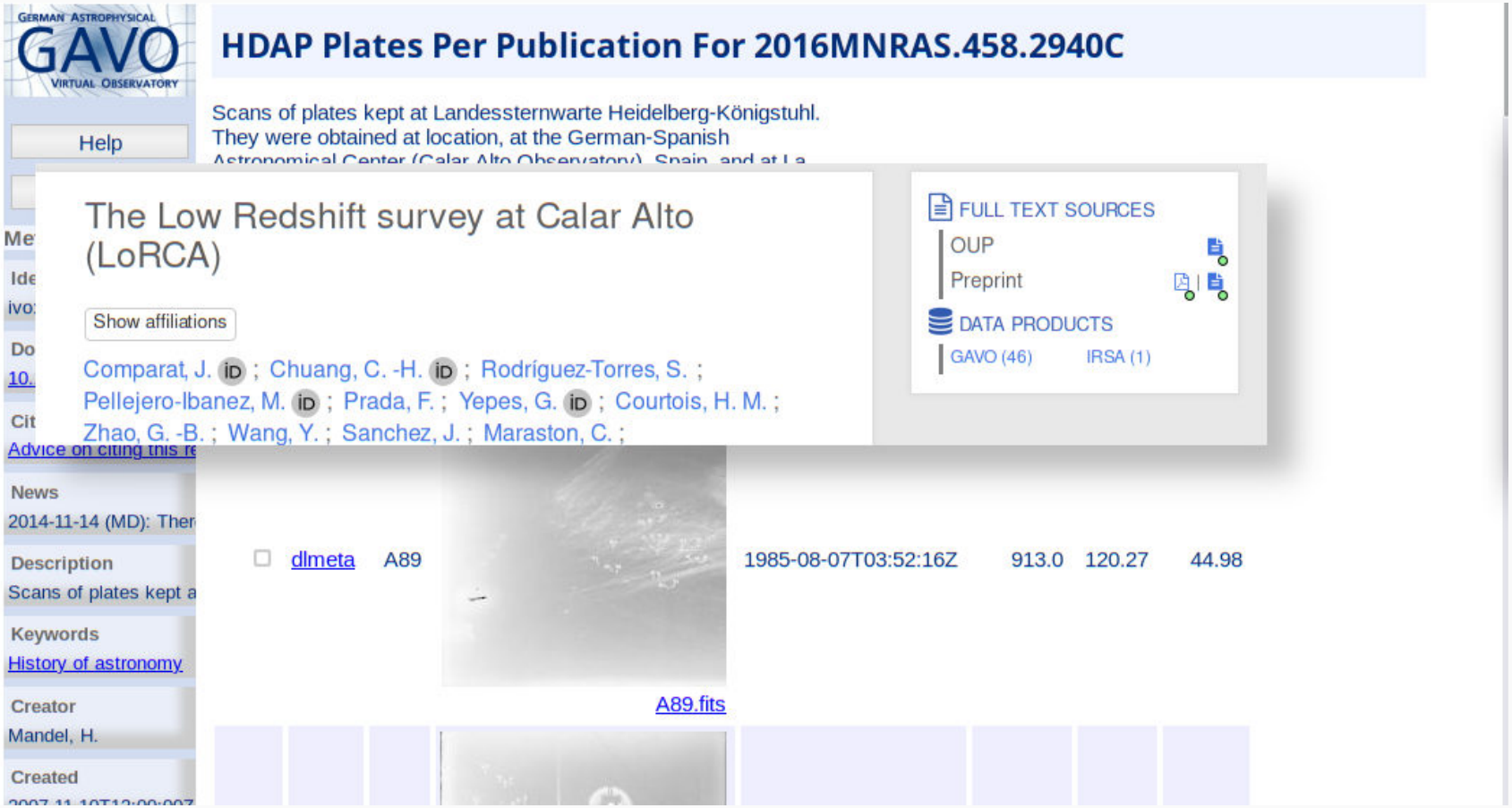
(And not at all if you run DaCHS).

biblink Landing Pages

SciX does not want to get 34 links from you when a paper has used 34 of your datasets. Instead, do a local landing page and tell them about the number of datasets:


```
{
  "bib-ref": "2019MNRAS.483.3773S",
  "relationship": "Cites",
  "dataset-ref": "http://[...]/biblanding/qp/2019MNRAS.483.3773S",
  "cardinality": 34
}
```

biblink in Reality



VO Resources as Bibliographic Records

ADS has been creating records for VizieR for a long time.

To give that option to other publishers having resources not directly linked to a paper, BibVO proposes:

```
<curation>
  [...]
  <date role="ExportRequested">2023-10-27</date>
</curation>
```

Then obtain a DOI (so SciX can consume DataCite rather than VOResource), and you would get a bibcode *for your VO resource.*

Disclaimer: SciX does not yet do that.

Conclusions

- Make sure your content/source is a resolvable bibcode (and the right one!)
- If you know datasets from your resources, tell SciX about it using biblink-harvest
- There are ~ 4 orders of magnitude more requests to SciX than to the Registry: It's worth it in terms of discovery!
- If you feed SciX, make sure that your infrastructure works: *Nobody* likes broken links from SciX