

2024-11-16 IVOA DCP Session

Landing page: <https://wiki.ivoa.net/twiki/bin/view/IVOA/InterOpNov2024DCP>

Agenda

Speaker	Title	Time	Materials	Abstract
G.Landais	Data Origin status	3	pdf	
P.Fernique	Core Trust Seal certification at CDS the HiPS extension	15+5	pdf	
Wu Xiuqin	What NED has done in telescope, instrument, bandpass, and frequency	10+5	pdf	

G.Landais	Assigning DOI for all VizieR Catalogues	10+5	pdf	The VizieR catalogue is harvested by network search as the VOREigstry or ADS. Dataset are also exposed through other Science network when DOI exists. In the presentation, I will outline our plan to extend the DOI generation to all VizieR .
M.Demleitner	Bibliographic Interfaces in the Virtual Observatory	15+5	notes slides	

all	dicussions (subject to define: DOI, Biblink, ...)	20	pdf	
-----	--	----	-----	--

Gilles on Data Origin

<INFO> Tags for Data Origin.

<INFO name='related_resource'>

Use in astroquery? To extract citation, references.

Question: Baptiste. Where is the IVOA id?

Question: Baptiste. Why URL and DOI? Use HTTP DOI for landing page.

Question: Anne — <https://doi.org/> is rec reference rather than the protocol.

Gus's reply. Please note that the request by DOI for <https://doi.org/> prefixed DOIs is a "display" request. According to the DOI handbook a DOI is still defined at the DOI string prefixed by doi: . The presentation request by DOI is that in "print" versions it is important to display the DOI as "actionaable" and to use a convention such as the https://doi.org prefix to show it as actionable. <https://www.doi.org/doi-handbook/HTML/presentation-formats-of-a-doi-.html>

Question: Tom D. — what is this DOI? Is it the source DOI, but not the DOI of the result. Is there a connection to which element the Data Origin INFO and DOI? Where can it appear?

Question: Markus — how to write DOIs. Bake in HTTPS is stupid (I agree). Big upstream "fix" necessary to switch to DOI.

Question: Anne — workarounds? Way tooo backwards.

```

<INFO name='protocol' value='Simple Cone Search 1.03'/>
<INFO name='request_date' value='2023-10-31T15:47:25'/>
<INFO name='request' value='https://vizier.cds.unistra.fr/viz-bin/conesearch/'>
<INFO name='contact' value='cds-question@unistra.fr'/>
<INFO name='version' value='7.32.5'/>

<INFO name='publisher' value='CDS'/>
<INFO name='creator' value='Ou J.-Y.'/>
<INFO name='related_resource' value='bibcode:2022AJ...163..1920'/>
<INFO name='landing_page' value='https://cdsarc.cds.unistra.fr/viz-bin/cat/J/AJ/163/192'/>
<INFO name='publication_id' value='doi:10.26093/cds/vizier.51630192'/>
<INFO name='publication_date' value='2022-09-26'/>

```

Aladain/CDS evaluating CTS.

Third time they do this exercise.

Big exercise. Finicky writing requirements. 3000eu cost!

- Be recognized as a reliable data repository



- Evaluate/modify our practices to improve our data management



Interop IVOA – Malta – Nov 2024 – P.Fernique

Information(KISTI)			Daegu, 34141, Korea, Daegu, KR	
Cambridge Structural Database	Cambridge Crystallographic Data Centre	https://www.ccdc.cam.ac.uk/	12 Union Road, Cambridge, CB2 1EZ, United Kingdom, Cambridge, GB	2027-03-28
GESIS Data Services	GESIS - Leibniz Institute for the Social Sciences	https://www.gesis.org/en/data-services/share-data	Quadrat B6 4-5, 68159 Mannheim, DE	2027-03-28
Global Hydrometeorology Resource Center	NASA and University of Alabama in Huntsville	https://ghrc.nasr.nasa.gov/home/	320 Sparkman Dr., Huntsville, US	2027-03-14
LINDAT-CLARIAH-CZ	Institute of Formal and Applied Linguistics	https://lindat.cz	Malostranské náměstí 25, Praha, CZ	2027-03-11
Strasbourg Astronomical Data Centre	Strasbourg Astronomical Observatory (UMR 7550, a Joint Unit of the Centre National de la Recherche Scientifique (CNRS) and of Strasbourg University)	https://cds.unistra.fr	11 rue de l'université, Strasbourg, FR	2027-03-11
PORTULAN CLARIN	Faculdade de Ciências da Universidade de Lisboa	https://portulan.clarin.net	Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa, Edifício C6, Campo Grande, 1748-016 Lisboa, Lisboa, PT	2027-03-11
RESIF seismological Data Repository	RESIF/CNRS/Université Grenoble Alpes	seismology.resif.fr	OSUG, 122 Rue de la Piscine, Saint Martin d'Hères, 38400, FR	2027-03-11
Tethys RDR	GeoSphere Austria	https://www.tethys.at/	Hohe Warte 38, 1190 Wien, Vienna, AT	2027-03-11
Physical Oceanography				

- 3rd CTS request for CDS
 1. DSA: 2014 , first certification (other acronym)
 2. CTS: 2019, second certification ended in 2022
 3. Next certification for 3 years more
- So far, only data from the Catalog Service (VizieR)
=> Extend this new application to all CDS data (10%
-> 99%) by certifying also HiPS images (Sky Atlas
Service - Aladin)

Wu NED's Fac List

Questions: Baptiste — Semantics working on vocab for terms. Relations for telescopes and facilities.



Facility/Observatory, Telescope, Instrument, Filter

Xiuqin Wu
NED task lead, Caltech/IPAC

Ivoa interop, Nov. 15-17, 2024, Malta

1

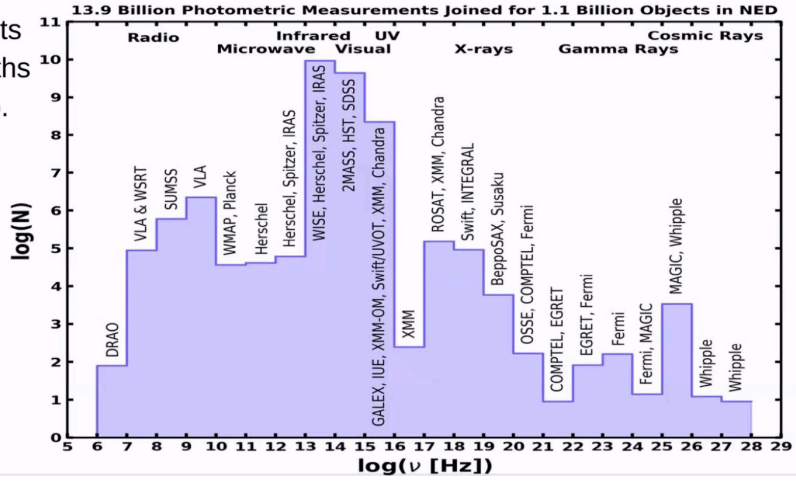
Standardizing photometry data points.



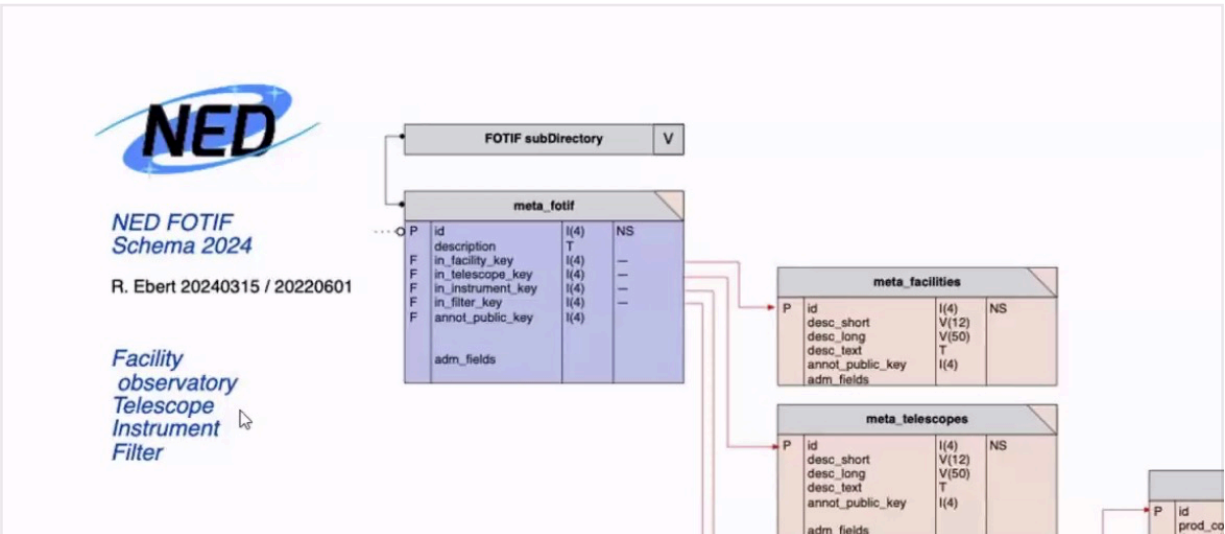
NED Photometry Data



- 13.9B data points
- Wide wavelengths
- 19K header info.
- Not consistent
 - WISE w1,
 - w1 WISE ...



Facility / observatory
 Telescope
 Instrument (they do not collect aliases)
 Filter



select * from meta_fotif_view limit 20;

id	facility	telescope	instrument	filter	freq_hi	freq_lo	freq_eff	ref_frame	ivoa_filter_id
20043	Arecibo, Green Bank, MPIFR	Arecibo, 91m, Effelsberg, Nancay					1.420e+09		
42625	Arecibo Obs	Arecibo	ALFA				1.420e+09		
47537	ASTRON	WSRT	Drift scan mode				1.420e+09		
24767	ATNF	Parkes					1.420e+09		
10304	ATNF	Parkes					1.420e+09		
48519	Chandra X-Ray Obs	Chandra	LETGS				1.150e+18		
15618	Green Bank Obs	GBT					1.420e+09		
48407	Jodrell Bank Obs	Lovell 76m	Multibeam receiver				1.420e+09		
33953	Mauna Kea Obs, CFH Obs, Austr. Astro. Obs	UKIRT, CFHT, AAT	IRCAM1, MONICA, IRIS					1.390e+14	
25259	Nancay Radio Obs	NRT					1.420e+09		
30210	NAOC Xinglong Obs	EST, TNT	CCD				4.680e+14		
44111	Nobeyama Radio Obs	NMA	RAINBOW interferometer					8.710e+10	
34991	NRAO	EVLA	D configuration				1.150e+11		
34573	Observatorio del Roque de los Muchachos	WHT	ISIS				4.690e+14		
30667	Observatorio do Pico dos Dias	1.6m	B&C spectrograph				4.450e+14		
41758	Palomar Obs	1.5m, 5m	Photomultiplier				5.421e+14		
8641	Smithsonian Astrophysical Obs/ASIAA	SMA					2.310e+11		
14172	Smithsonian Astrophysical Obs/ASIAA	SMA					2.310e+11		
33048	Spitzer	Spitzer	IRS				3.840e+13		
36184	University of Hawaii	2.2m	Tek2k				4.640e+14		

From Baptiste's semantics talk this afternoon.

ObsFacility Vocabulary

Almost there!

- Long standing action: gathering a list of terms for observation facilities.
- Required for interoperability and discoverability of dataset based on telescope name or spacecraft name.
- Semantic reminder:
Observation Facility = the concept bearing the location of the observation
Instrument = the concept bearing the instrumental capabilities (fov, filter...)
- Current scope:
 - a reference list of terms for interoperability.
 - a name resolver to help finding the right term to be used.
 - hierarchy (has_part, is_part_of)
- Not in scope:
 - modelling the observation facility (classification, metadata, relation to instruments)
 - instruments

ObsFacility Vocabulary

Almost there!

- Sources: SPASE, PDS, ObsCode, NSSDC, NAIF, AAS, etc, and Wikidata
Wikidata has a lot of identifiers already mapped together: great starting point.
But very difficult to curate not being part of wikidata community.
- Process:
 - Export from Wikidata and merge additional terms and aliases and produce a curated list of terms with aliases.
 - list of terms + external identifiers + relations => IVOA Vocabulary
 - list of terms + known aliases / ids => Name resolver
- Plan:
 - first version with terms having an ObsCode or a SPASE id or a PDS id (almost there)
 - decisions to be made on the accepted terms, curation of label and description
 - second version making sure that all VOResource/ObsCore/EPNcore observation facility values can be mapped to a term in the vocabulary.
 - then: updates through VEP.

ObsFacility Vocabulary

Example

```
obs:gemin-north a owl:Class ;
  rdfs:label "Gemini North"@en ;
  rdfs:comment "northern hemisphere facility of the Gemini Observatory (wikidata:Q6148627) (obscode:T15)"@en ;
  skos:altLabel "Q6148627",
    "T15",
    "frederick c. gillett gemini telescope"@en,
    "gemini north observatory"@en,
    "gemini north telescope"@en,
    "the frederick c. gillett gemini telescope"@en ;
  skos:exactMatch <http://www.wikidata.org/entity/Q6148627>,
    <https://minorplanetcenter.net/iau/lists/ObsCodesF.html#T15>,
    <urn:nasa:pds:context:facility:observatory.gemini-north-maunakea> .

obs:gemin-south a owl:Class ;
  rdfs:label "Gemini South"@en ;
  rdfs:comment "southern hemisphere facility of the Gemini Observatory (wikidata:Q19673584) (obscode:I11)"@en ;
  skos:altLabel "I11",
    "Q19673584",
    "gemini south observatory"@en,
    "gemini south telescope"@en,
    "gems"@en ;
  skos:exactMatch <http://www.wikidata.org/entity/Q19673584>,
    <https://minorplanetcenter.net/iau/lists/ObsCodesF.html#I11>,
    <urn:nasa:pds:context:facility:observatory.gemini-south> .
```

Gilles on DOI in Vizier

1 Vizier Catalogue for ALL the files (collection)

And 1 DOI for that catalogue/collection.

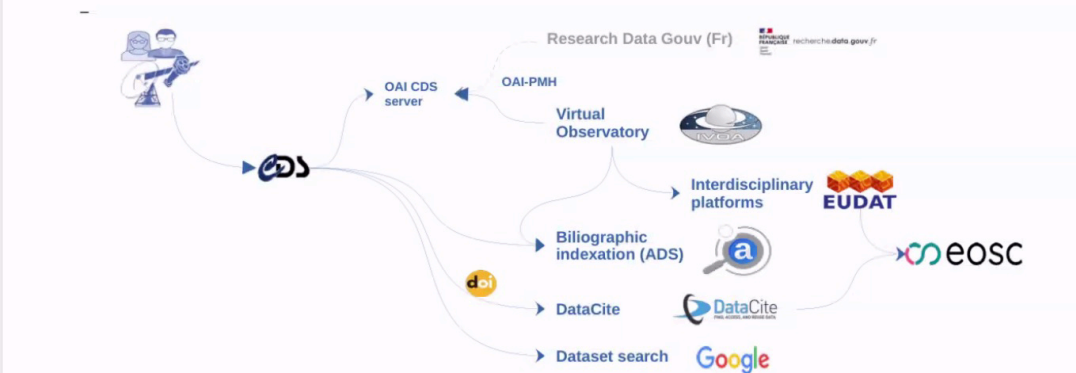
Broadcast authors datasets to a wide audience

Granularity is the Catalogue

ie: collection of files : FITS and tables curated y CDS

- coming from an author article (ie: authors submission)
- coming from space agencies or ground based telescop

Providing metadata using the appropriate schema



Gus: I like that workflow, Gilles!

Linked Data: more granularity?

EOSC uses relations but ignores ones that are not DOI.

Description is the abstract of article. <no!>

Gus: Not the same bc abstract doesn't describe the data does it?

Baptiste: Idea that we could add identifiers that are reference IDs that could be accepted on their graph w/o DOIs.

- Gus: Bibcodes are going away anyway.

□ Dissemination



The same granularity in all networks, but ...

- Same metadata for : authors, DOI, Title, dates, keywords (UAT)
- Specific metadata
 - Linked Data to service (only Registry, B2Find)
 - Linked Data (related identifier) specific (Registry) or not used (EOSC)
 - ORCID (only in DataCite)
 - Specific Descriptions : abstract (Registry/Eudat), specific (DataCite/Eosc)



Status of the DOI workflow

- Automated workflow for every catalogue coming from journals : AAS (AJ, ApJ, ApJS), A&A, MNRAS
~80% of vizier catalogue have DOI
- DOIs curated by “hand” for ESA catalogues
 eg: Gaia DR3: doi:10.26093/cds/vizier.1355

EUDAT: links to resources

EOSC: found some other information (funding) How?

Mass-metallicity relation for giant planets (Kong S. et al., 2018)
<https://cdsarc.cds.unistra.fr/viz-bin/cat/J/ApJ/831/64>

Mass-metallicity relation for giant planets

doi
 Exoplanet discoveries of recent years have provided a great deal of new data for studying the bulk compositions of giant planets. Here we identify 47 transiting giant planets ($20 M_{\oplus} < M < 20 M_{\text{J}}$) whose stellar insulations are low enough ($F_{\text{inc}} < 2 \times 10^8 \text{ erg s}^{-1} \text{ cm}^{-2}$, or roughly $T_{\text{eff}} 50 M_{\oplus}(\text{Earth})$) suggest significant amounts of heavy elements in H/He envelopes, rather than cores, such that metal-enriched giant planet atmospheres should be the rule. We also discuss a model of core-accretion planet formation in a one-dimensional disk and show that it agrees well with our derived relation between mass and $Z_{\text{planet}}/Z_{\text{star}}$.

Cone search capability for table JApJ831/64table1 (Planet data and results)

chemical abundances | interdisciplinary a... | solar system astronomy | solar system planets
 stellar astronomy | stellar masses

Identifier	
DOI	http://doi.org/10.26093/cds/vizier.18310064
Source	https://cds-vo.org/rfq/lp/custom/CDS.VizieR/JApJ/831/64
Related Identifier	https://cdsarc.cds.unistra.fr/viz-bin/cat/JApJ/831/64
Related Identifier	http://vizier.cds.unistra.fr/viz-bin/VizieR-2?-source=JApJ/831/64

DATA DATASET

Year: 2017 | Views: 0 | Downloads: 0 | Citations: 0

Mass-metallicity relation for giant planets

VizieR online Data Catalogue associated with article published in journal *Astronomical Journal* (AAS) with title 'The mass-metallicity relation for giant planets.' (bibcode: 2016ApJ...831...64T)

Author Thomgren, D.P. | Fortney, J.J. | Murray-Clay, R.A. | Lopez, E.D.

Publisher Strasbourg Astronomical Data Center | B2FIND

Identifier 10.26093/cds/vizier.18310064 (DOI)

Funding <https://cordis.europa.eu/project/id/313014> (European Commission)

Keywords stellar astronomy • Chemical abundances • Solar system planets • Solar System Astronomy • solar system astronomy • interdisciplinary...

IVOA 2024, Malta - DOI status for VizieR catalogue

isSupplementTo ... YCATS?? No, please! A YCAT is an expression of the Vizier data.

CDS Proposals based on existing VizieR DOI

- UAT keywords
- Title and description adapted to the dataset
- Generic rights (to improve)
- Linked Data

VizieR online Data Catalogue associated with article published in journal *Astronomical Journal* (AAS) with title 'WYIN open cluster study. VII. NGC 2451A and the Hipparcos distance scale.' (bibcode: 2001AJ....122.1486P)

- Alternate identifier to link VO
- Link `IsSupplementTo` to the article for dataset coming from space agencies
- Link `IsVariantFormOf` the original dataset
- Cite `cites` original Data

```
"relatedIdentifiers": [  
  {  
    "relationType": "IsSupplementTo",  
    "relatedIdentifier": "2022yCat.1356....0G",  
    "relatedIdentifierType": "bibcode"  
  },  
  {  
    "relationType": "IsVariantFormOf",  
    "relatedIdentifier": "10.5270/esa-qa4lep3",  
    "relatedIdentifierType": "DOI"  
  },  
  {  
    "relationType": "Cites",  
    "relatedIdentifier": "10.5270/esa-qa4lep3",  
    "relatedIdentifierType": "DOI"  
  }  
]
```

IVOA 2024, Malta - DOI status for VizieR catalogue

20 % VizieR catalogues with no DOI

Question of creating CDS DOI is asked to data producers (editors, space agencies, institutes) who have data in vizier



Markus on BibVO

Three sections in BibVO.

Gus: there's directionality in those relationships, If ADS is using it then ADS's record already uses.

I. Data for an article.

=> <content><source> tag for the bibcode. Or DOI. Or Free text.

=>

Shows up as a "D-link"

ADS has IRSA data links, but Markus's isn't sure where that comes from.

II. BibVO for multiple objects. Main part is dealing with these. These could be created or used objects.

ADS does have this linking capability to support observatory bibliographies.

biblink-harvest.

- 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).
 - ADS does not want 34 granular things to use.
 - ADS wants a summary of the # of objects (cardinality)
 - GAVO (34).
 - AAS (15)
- ADS harvests and updates links weekly.

Open discussions



- <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?

III. Publishing VOResources specifically when the VOResource has no publication record elsewhere — to get them into ADS (whenever ADS indexes datasets)


<curation> tags

VOResource DOIs as proxy for VOResource metadata.

ADS doesn't do that yet.



GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung



Bibliographic Interfaces in the VO

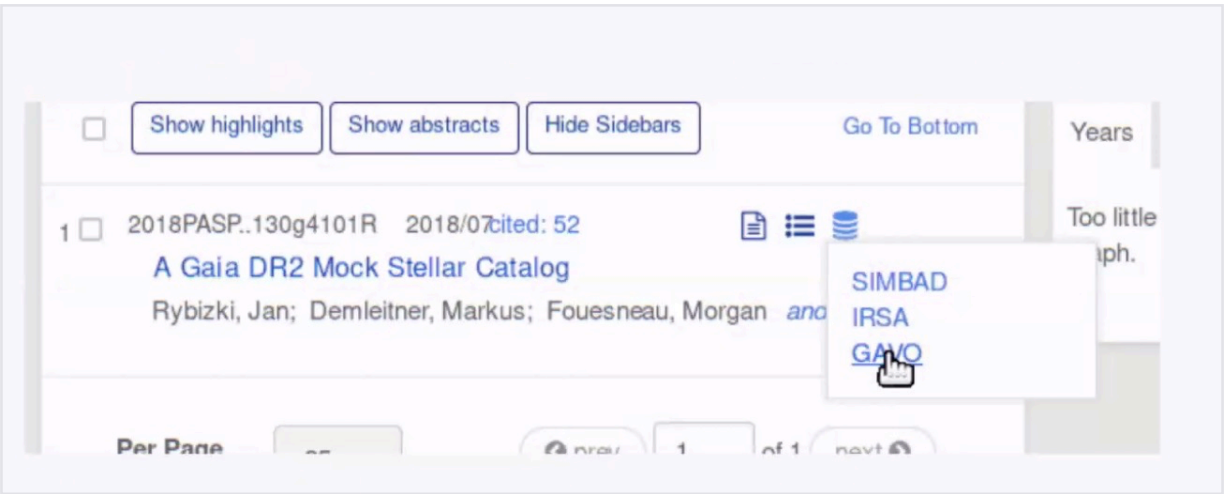
Markus Demleitner
IVOA Southern Spring Interop 2024, Malta Nov 14-17
Data Curation and Preservation IG

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.



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`

- 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