

# IVOA Registry/DCP Session: May 2023 Bologna

Schedule: <https://wiki.ivoa.net/twiki/bin/view/IVOA/InterOpMay2023RegistryDCP>

Plenary Room: about 50 attendees  
+ 22 online participants

## **S.Derriere - Description of HiPS surveys in the IVOA registry**

Initiated at CDS

HiPS: IVOA std for publishing image surveys, also catalogues

20 servers incl. 3 at CDS, 1200 individual HiPS

HiPS metadata reside in properties ASCII file with key=value lines

CDS pub. registry with OAI endpoint was upgraded in 2020 to use pyaoi, contains >24k resources (+1200/year). It contains mostly Vizier catalogues as vs:CatalogServices and 3 reg records for HiPS servers. In 2023 there was no desc of individual HiPS so they are going to be added define mapping...

Individual HiPS registration is described in section 5.3 of the HiPS std using the hips key

Best interface type ? ParamHTTP, dir ? HiPS files are in a directory, but there is a HTML landing page so 'dir' no right.

HiPS catalogs: some v large catalogs have progressive HiPS access in addition to Web, ParamHTTP/VOT, SCS, TAP. Example: Gaia DR3 part 1. 52 HiPS catalogue capabilities have been added, +1 for SIMBAD

HiPS surveys: use UAT term for <subject>

Mapping from properties file to VOResource

Conclusion: 543 HiPS in the CDS Registry with dedicated OAI-PMH set: "HiPS", also visible in ???

Perspectives: Add DOIs to Resources. How to search for individual HiPS in the VO: Using RegTAP=>add Template queries to Search Interfaces (similar to SCS, SIA, or TAP)

Q&A

- AN: 2 separate registries for images and catalogs ? yes

## **G.Landais - Data Origin in the VO**

Splinter meeting in Oct2022 => IVOA Note

Data Origin? (DO) = info on origin of distributed data from selections, changes in dataflow or in a user query

Idea: Decorate VO results with metadata describing DO

Metadata in result: Dublin Core(id, authors, license) and other for access (data center, URL, params, date of execution)

Motivation: improve data understanding for end users, reproducibility and citation

Where to find DO in the VO: Registry, (DublinCore+Datacite compatible) DMs, VOT (with MIVOT later) Protocols ?

Metadata proposal: note in repo `ivoa-scp-data-origin` on github:

1 List of metadata for reproducibility and 1 list for tracing origin

Simple VOTable serialization based on INFO tag, well integrated in TOPCAT client + fork of Astropy by MD. Also available in Aladin

Acknowledgment extraction for automatic citation using a template

Conclusion a light provenance mechanism well integrated in VO framework, implemented in Vizier (beta) for SCS and ASU access. How to add that in DALI Q&A

- H.Enke about Semantics of metadata ? ... start an issue on the repo

- PLS: include DOI in metadata proposal ?

## **M.Demleitner - DOIs for everyone in the VO: ! VOiDOI**

ivoid vs DOI

Register your VO service => get a ivoid. which can be resolved using GAVO resolver <http://dc.go-vo.org/I/<ivoid>>

To generate a DOI from registry record, you can use <http://dc.g-vo.org/voidoi/q/ui>

You'll get an email at you registered contact address. WARNING: DOIs are for eternity. they don't go away !

You'll get a generated landing page that points you to a reformatted Reg Record or a "tombstone"

Q&A:

- Mark Allen: will the landing page be provided forever ? What is service disappear ? DOIs have a cost

- Tom: why have a DOI for something so transient as a VO Resource ? the landing page is useful to provide links access through VO protocols

- Christophe A: case with many services?

## **H.Enke (remote) - Using DOI on astronomical data @ AIP**

DOIs applied @AIP since 2011 for tables of RAVE DR3

DOI minting principles @AIP: use DOI as identifier for published data, not as data collection identifiers. DOI structure should convey info about collection. DOI

Metadata from IVOA data, landing page

Problems: v rapid dev of metadata scheme: DataCite Kernel: 2.x in 2011 to 4.x in 2023, 5.x in preparation ! This makes maintenance of data collection DOI difficult for the data provider.

Benefits: improved means to connect data collection to published pages

Managing Metadata for IVOA and DataCite: **Daiquiri** = django (+celery) based data publication framework

Using Datalink to mix info from webapp and science databases. Example page

Conclusion: considerable overlap between IVOA and DataCite metadata

## **Y.Tao - the NADC implementation of DOI**

NADC Data Resources: observation datasets from local telescopes LAMOST, FAST, BASS... + mirror of Gaia, SDSS, DESI... + Research Project data

Different PIDs are maintained according to data types

Granularity is defined case by case according to dataset peculiarities

DOIs for paper-related data are assigned on request

DOI vs CSTR=China Science and Tech Resource Identifier

Access and Citation Tracking

Conclusion: DOI = useful PID adopted by NADC

Q&A

- H.Enke: what is done at DOI level when metadata change ? same as Zenodo...

Continue this discussion on slack ?

## **R.D'Abrusco (remote) - A Chandra-centric approach to DOIs**

CDA/Chandra DOI strategy: used for

- Archival Observations DOIs(ObsID based)

- Chandra Source Catalogs and

- Chandra Data Collection (CDC)

Different types of aggregations of Chandra data are possible

Using RelationTypes to show relationships for RelatedIdentifiers brings a network

How to allow citability of subsets ?

## **Christophe Arviset - ESA Data Discovery Portal, link to ESA datasets DOIs and to Google Dataset Search**

Science Data Archives at ESA: Planetary Science Archive, Earth Science, etc.

DOIs have been defined for every ESA dataset using a common set of metadata, first with CrossRef and now with DataCite

Granularity level for DOI assignment in space science depends on the type of mission

- for Surveys (Gaia, Planck...) a few DOIs required

- for Observatory (HST, XMM...) a few 1000s of DOIs linked to scientific proposal

- for Planetary missions, DOIs are organized by datasets for instruments a few 1000s DOIs too

- for HelioPhysics, 1 DOI for each instrument/experiment thr the whole mission time

Discovery Portal data.esa.int built on top of the DOI database, provides a simple search interface. It's not an actual data archive, it only gather common metadata and does not replace the existing mission data archives

**Google Dataset Search:** datasetsearch.research.google.com released in 2018 targeting Open Science, references 25M datasets referenced thr schema.org

Q&A;

- PF re How to manage granularity in search ?