



---

## DOI into VizieR



Contributors:

Engineers: G.Landais, S.Wendell, A.Schaaff

Documentalists: E.Perret, M.Brouty ,

T.Pouvreau, P.Vannier

Astronomers: P.Ocvirk



And also: M.Allen, P.A.Duc, S.Langenbacher



CENTRE DE DONNÉES  
ASTRONOMIQUES DE STRASBOURG

# Identifiers in VizieR



## Identifiers provided by VizieR

- The catalogue name :  
example : I/345 (Gaia DR2)  
J/A+A/414/699 (catalogue A&A)
- The bibcode : shared with ADS  
2018A&A...616A...1G (Gaia DR2)  
2004A&A...414..699C (catalogue A&A)
- The ivoid (VO-registry)  
ivo://cds.vizier/i/345 (Gaia DR2)  
ivo://cds.vizier/j/a+a/414/699 (catalogue A&A)



## External identifiers (available in VizieR)

(added recently in VizieR metadata and updated from ADS database)

- Article DOI
- First author Orcid



# □ To generate DOI



- DOI registration organisation : Datacite  
<https://www.datacite.org/>



Only members can create DOI (in France, INIST (CNRS))

INIST: Institut de l'Information Scientifique et Technique  
<http://www.inist.fr/?Attribution-de-DOI&lang=fr>



- Recommendation given by INIST (CNRS):
  - Define the type, workflows of the DOI targeted
  - Verify the relevance of the cited object (granularity level? does the object merit to be cited?)
  - Data & metadata quality:
    - list the metadata and verify if they are compliant with the ethics

# □ To generate DOI



## Some requirements

- Stable URL
- Long-term preservation
- Access to the objects through the "landing pages"
- *(INIST) Meta-data must be available and subject to a CC0 licence*  
=> no restrictions on reuse of the metadata





# □ Why a VizieR DOI ?



## To take advantage of DOI

- DOI is a standard : it is commonly used by authors to cite data in article
- Search engine (datacite: <https://www.datacite.org/>)
- The URL redirection mechanism to a “landing page”
- DataCite provides statistics : <https://stats.datacite.org/#tab-resolution-report>

Secure | <https://stats.datacite.org>

### DataCite Statistics

Support

Registrations by Allocators   Registrations by Datacentres   Registrations by Prefixes   Resolutions by Month

Allocator	DOI Registrations				Metadata			
	Total	2017	2018	Last 30 Days	Searchable	Hidden	Missing	Ratio
ANDS - Australian National Data Service (ANDS)	280 207	45 316	65 906	3 475	277 922	1 379	906	99%
AU - American University	37	23	5	1	37	0	0	100%
AUSTINTX - City of Austin, Texas	20	0	20	20	6	14	0	100%
BIBSYS - Unit - The Norwegian Directorate for ICT and Joint Services in Higher Education and Research	3 654	226	3 044	415	3 642	12	0	100%
BL - British Library	1 169 051	103 123	98 809	20 305	1 151 484	17 538	128	99%
BRAINL - Brain Life	128	0	128	15	128	0	0	100%
BROWN - Brown University Library	3 072	273	2 629	109	3 071	1	0	100%
CALTECH - California Institute of Technology (Caltech)	1 452	523	253	14	1 442	10	0	100%
CCOM - Center for Coastal & Ocean Mapping - University of New Hampshire	4	0	4	4	0	4	0	100%
CDL - California Digital Library	1 408 962	689 889	241 079	34 861	1 402 743	6 219	3	99%
CERN - CERN - European Organization for Nuclear Research	828 623	224 727	236 722	27 590	660 560	168 063	0	100%
CHOP - Children's Hospital of Philadelphia	1	0	1	0	1	0	0	100%

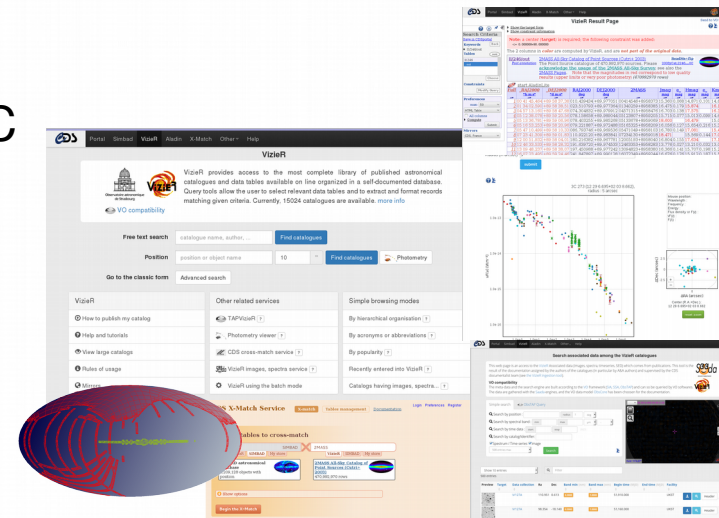
# □ What is a VizieR DOI ?



## What is VizieR DOI ?

**Granularity** : the catalogue (!= table in the IVOID registry)

- table can be subject to transformation even after publication (example : fusion)
- Include the whole catalogue :
  - The data : tables, associated data (FITS, spectra, images)
  - The metadata (ReadMe) : coordinates, filters, time, ...
  - Added values : links, plots....
  - The services to access the data :
    - VO : SCS, TAP, SIA, SSA, ObsTAP, MOC
    - FTP
    - Dedicated web application



# □ DOI metadata



**A join effort with the VizieR team** - documentalists + engineer

- The VizieR DOI syntax :  
DataCite encourages DOI publishers to use a sufficiently opaque suffix
  - example : XXXX/CDS/xxxxx  
Prefix  
(fixed by DataCite)      Suffix  
(specific to each object and fixed by VizieR)
  - DOI is persistent  
=> Opaque suffix is preferable to avoid information in the identifier  
  
e.g. : catalogue removed after being published in VizieR
- Define DOI metadata AND metadata in the landing page  
→ Privilege meta-data in the landing-page rather than DOI metadata
  - more flexible !
  - What are the benefits to put all metadata in DOI ?  
(example : abstract, ..)

- DOI based on VizieR catalogue name :  
XXXX/CDS/VizieR/J/A+A/1246
- Semi-opaque DOI :  
XXXX/CDS/34140699
- Opaque DOI :  
XXXX/34140699




# □ DOI metadata



- Datacite provides the MDS service (Metadata Store) to register DOI associated with metadata.
- DOI Metadata are gathered into a XML document :  
[http://schema.datacite.org/meta/kernel-4/doc/DataCite-MetadataKernel\\_v4.0.pdf](http://schema.datacite.org/meta/kernel-4/doc/DataCite-MetadataKernel_v4.0.pdf)

## Metadata subject to a particular attention in the VizieR context

- Remove abstract from DOI-metadata because abstract could be subject to licences 
- Related identifiers :
  - Alternative identifiers : **bibcode**, **IVOID**
  - Link external identifiers with the relationship type:
    - **DOI** : the Datacite XML schema provides a rich grammar :  
*IsCitedBy, IsSupplementTo, IsSupplementedBy, IsContinuedBy, IsPartOf, HasPart, IsReferencedBy, ...*

→ Use «*IsSupplementTo*» relation to link VizieR catalogues with published articles

```
<relatedIdentifier relatedIdentifierType="DOI" relationType="IsSupplementTo">  
  10.1051/0004-6361:20031671  
</relatedIdentifier>
```

- **Orcid** of the first author enables link (URL resolution similar than DOI)

→ populate DOI & Orcid in VizieR from **ADS**





# Comparison with registry



## IVOA registry

## DOI metadata

```
<?xml version="1.0" encoding="UTF-8" ?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:ri="http://www.ivoa.net/xml/RegistryInterface/v1.0"
  http://www.ivoa.net/xml/RegistryInterface/RegistryInterface-1.0.xsd">
  <responseDate>2018-10-31T17:13:24Z</responseDate>
  <request metadataPrefix="ivo_vor" verb="GetRecord" identifier="ivo://CDS.VizieR/J/A+A/414/699">
    http://cdsweb.u-strasbg.fr/reg-bin/vizier/oi_aux.pl
  </request>
  <GetRecord>
    <record>
      <header>
        <identifier>ivo://CDS.VizieR/J/A+A/414/699</identifier>
        <datestamp>2018-04-05T10:00:00Z</datestamp>
        <setSpec>J.A.A</setSpec>
        <setSpec>ivo_managed</setSpec>
      </header>
      <metadata>
        <ri:Resource xmlns="" xmlns:ri="http://www.ivoa.net/xml/RegistryInterface/v1.0" xmlns:vr="http://www.ivoa.net/xml/ResourceService"
          created="2017-06-19T07:57:41" updated="2018-04-05T10:00:00" status="active"
          xsi:type="vs:CatalogService" http://www.ivoa.net/xml/ConeSearch/v1.0">
          <validationLevel validatedBy="ivo://CDS.VizieR"></validationLevel>
          <title>Spectra of southern late-type dwarfs (Cincunegui, 2004)</title>
          <shortName>J/A+A/414/699</shortName>
          <identifier>ivo://CDS.VizieR/J/A+A/414/699</identifier>
          <curator>
            <publisher ivo-id="ivo://CDS">CDS</publisher>
            <creator>
              <name>Cincunegui C., Mauas P.J.D.</name>
            </creator>
            <contributor>Patricia Bauer [CDS]</contributor>
            <date role="creation">2017-06-19T07:57:41</date>
            <version>04-Nov-2003</version>
            <contact>
              <name>CDS support team</name>
              <address>CDS, Observatoire de Strasbourg, 11 rue de l'Universite, F-67000 Strasbourg, France</address>
              <email>cds-question@unistra.fr</email>
            </contact>
            <content>
              <subject>Spectroscopy</subject>
              <subject>Stars:late-type</subject>
              <description>We present Echelle spectra of 91 late-type dwarfs, of spectral types from F to M and of different (lambda)26400). The observations were flux calibrated with the aid of long slit spectra (R=1050).</description>
              <source format="bibcode">2004A&A...414..699C</source>
              <referenceURL>http://cdsarc.u-strasbg.fr/cgi-bin/Cat?J/A+A/414/699</referenceURL>
              <type>Catalog</type>
              <contentLevel>Research</contentLevel>
              <relationship>
                <relationshipType>IsServedBy</relationshipType>
                <relatedResource ivo-id="ivo://CDS.VizieR/TAP">TAP VizieR generic service</relatedResource>
              </relationship>
              <rights>public</rights>
              <capability>...</capability>
              <capability standardID="ivo://ivoa.net/std/TAP#aux">...</capability>
              <capability xsi:type="cs:ConeSearch" standardID="ivo://ivoa.net/std/ConeSearch">...</capability>
              <coverage>
                <footprint ivo-id="ivo://mocivod">http://alaska.u-s...</footprint>
                <waveband>Optical</waveband>
              </coverage>
              <table>
                <name>J/A+A/414/699</name>
                <description>Observ...</description>
                <column>...</column>
                <column>...</column>
                <column>...</column>
              </table>
            </ri:Resource>
          </metadata>
        </record>
      </GetRecord>
    </OAI-PMH>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<resource xsi:schemaLocation="http://datacite.org/schema/kernel-4 http://schema.datacite.org/meta/kernel-4.1/metadata.xsd">
  <identifier identifierType="DOI">YYYYY/CDS/VizieR.1246</identifier>
  <creators>
    <creator>
      <creatorName>Cincunegui C.</creatorName>
    </creator>
    <creator>
      <creatorName> Mauas P.J.D.</creatorName>
    </creator>
  </creators>
  <titles>
    <title xml:lang="en">Spectra of southern late-type dwarfs</title>
  </titles>
  <publisher>Centre de Donnees Strasbourg (CDS)</publisher>
  <publicationYear>2004</publicationYear>
  <subjects>
    <subject schemeURI="http://cdsarc.u-strasbg.fr/doc/ADCKwds.htx" subjectScheme="ADC Keywords"> Atlases </subject>
    <subject schemeURI="http://cdsarc.u-strasbg.fr/doc/ADCKwds.htx" subjectScheme="ADC Keywords"> Spectroscopy </subject>
    <subject schemeURI="http://cdsarc.u-strasbg.fr/doc/ADCKwds.htx" subjectScheme="ADC Keywords"> Stars, late-type </subject>
  </subjects>
  <contributors>
    <contributor contributorType="DataCurator">
      <contributorName>CDS team</contributorName>
    </contributor>
  </contributors>
  <dates>
    <date dateType="Created">22-Jan-2004</date>
    <date dateType="Updated">19-Jun-2017</date>
  </dates>
  <resourceType resourceTypeGeneral="Dataset">Dataset</resourceType>
  <alternateIdentifiers>
    <alternateIdentifier alternateIdentifierType="internal ID">J/A+A/414/699</alternateIdentifier>
    <alternateIdentifier alternateIdentifierType="ivoid">ivo://CDS.VizieR/j/a+a/414/699</alternateIdentifier>
  </alternateIdentifiers>
  <relatedIdentifiers>
    <relatedIdentifier relatedIdentifierType="bibcode" relationType="IsSupplementTo">2</relatedIdentifier>
    <relatedIdentifier relatedIdentifierType="DOI" relationType="IsSupplementTo">10.1051/0004-6361:20031671</relatedIdentifier>
  </relatedIdentifiers>
  <rightsList>
    <rights rightsURI="http://cds.u-strasbg.fr/vizier-org/licences_vizier.html">Refer to CDS usage</rights>
  </rightsList>
  <geoLocations>
    <geoLocationPlace>Strasbourg astronomical Observatory, France</geoLocationPlace>
  </geoLocations>
</resource>
```

### Metadata missing compared to DOI

- Identifiers e.g: ivoid, bibcode, orcid
- Rights : (url to a general web page)

### Metadata missing compared to registry

- Limited content description:
  - tables and columns
- Links : the landing page is given with metadata (but outside the XML)



## Spectra of southern late-type dwarfs : [J/A+A/414/699](#)

Access to

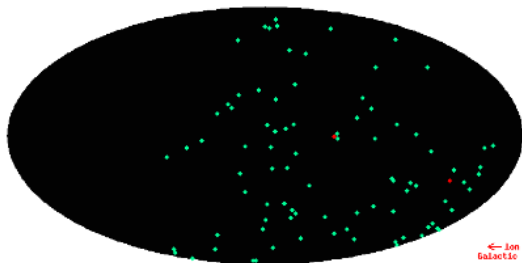


Authors : Cincunegui C. , Mauas P.J.D.

**Bibcode** : 2004A&A...414..699C (ADS)  
2004A&A...414..699C (Simbad)  
[Simbad objects](#)

**CDS Keywords** : Atlases ; Spectroscopy; Stars, late-type

Observation (OC)



Inserted into VizieR : 22-Jan-2004  
Last modification : 19-Jun-2017

Article Origin [Acknowledgment](#) [ReadMe](#) [FTP](#) **[VizieR](#)**

Library of flux-calibrated echelle spectra of southern late-type dwarfs with different activity levels.

Article DOI : 10.1051/0004-6361:20031671

Keywords : atlases - stars: late-type

**Abstract**: We present Echelle spectra of 91 late-type dwarfs, of spectral types from F to M and of different levels of chromospheric activity, obtained with the 2.15-m telescope of the CASLEO Observatory located in the Argentinean Andes. Our observations range from 3890 to 6690Å, at a spectral resolution from 0.141 to 0.249Å per pixel ( $R=\lambda/\delta\lambda\sim 26400$ ). The observations were flux calibrated with the aid of long slit spectra ( $R\sim 1050-2070$ ), also available.

<http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/414/699>

# □ Landing page



## Gaia DR2 : I/345

Access to



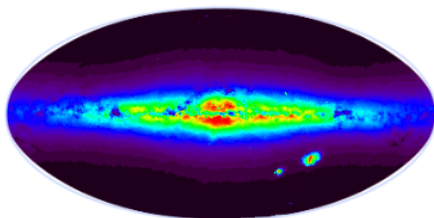
Authors : Gaia collaboration

Bibcode : 2018A&A...616A...1G (ADS)  
2018A&A...616A...1G (Simbad)  
Simbad objects

CDS Keywords : Surveys ; Stars, standard ; Positional data ; Proper motions ; Photometry, photographic ; Cross identifications ; Radial velocities ; Stars, variable ; Minor planets; Parallaxes, trigonometric

Observation (OC)

Records : 1692919135 sources



Inserted into VizieR : 16-Mar-2018  
Last modification : 19-Oct-2018

Article Origin Description Acknowledgment See also ReadMe FTP **VizieR**

Gaia data release 2 (Gaia DR2).

Article DOI : [10.1051/0004-6361/201833051](https://doi.org/10.1051/0004-6361/201833051)

Keywords : catalogs - astrometry - techniques: radial velocities - stars: fundamental parameters - stars: variables: general - minor planets, asteroids: general

**Abstract:** Gaia Data Release 2. Summary of the contents and survey properties: We present the second Gaia data release, Gaia DR2, consisting of astrometry, photometry, radial velocities, and information on astrophysical parameters and variability, for sources brighter than magnitude 21. In addition epoch astrometry and photometry are provided for a modest sample of minor planets in the solar system. A summary of the contents of Gaia DR2 is presented, accompanied by a discussion on the differences with respect to Gaia DR1 and an overview of the main limitations which are still present in the survey. Recommendations are made on the responsible use of Gaia DR2 results. ...[\(more\)](#)

<http://cdsarc.u-strasbg.fr/viz-bin/cat/I/345>

Go to the Gaia DR2 portal





## Finalise the VizieR DOI workflow

- Signature of the contract with INIST (CNRS)
- Include the DOI workflow into the VizieR ingestion workflow
  - DataCite provide a test environment (available for members)
    - A web page (requiring authentication)
    - MDS API : REST API <https://mds.datacite.org>  
(requires organisations to first register for an account with a DataCite member)
- Generate our first DOI : a DOI for VizieR
- DOI Generation won't be automated in a first step..
- Take into account the sustainability of DOI
  - landing page must be sustainable .
  - create a generic deprecated VizieR page for removed catalogue (after publication)



The screenshot shows the 'Enregistrer nouveau Ensemble de données' (Register new data set) form in the DataCite Metadata Store. The form includes a sidebar with navigation options like 'Enregistrer nouveau/nouvelle Ensemble de données', 'Énumérer tous/toutes les Ensembles de données', and 'Rechercher par DOI'. The main form area contains a warning about DOI latency, a field for 'DOI' (circled in blue), a field for 'Adresse URL' (circled in blue), and a file upload section for 'Téléversement d'entités XML' with a 'Parcourir...' button (circled in blue). Below these fields is a note about choosing an XML file and a base URL, followed by an 'XML:' text area and an 'ENREGISTRER' button.