

# Challenges and roles in electronic data publication

P. Ocvirk, M. Allen  
CDS, Observatoire de Strasbourg



- 1) Vizier
- 2) Challenges, the paper legacy
- 3) Workflow
- 4) Roles



VizieR Staff and contributors:

Astronomers: P. Ocvirk, L. Cambresy, F. Ochsenbein, F. Genova, S. Derriere

Engineers: G. Landais, T. Boch, F.X. Pineau, A.C. Simon, F. Bonnarel

Documentalists: P. Vannier, E. Perret, M. Brouty, S. Guéhenneux

External: L. Michel, J.Y. Hangouet, T. Keller (Strasbourg Observatory)

Vizier is a **service** which gives a unified access to a very large collection of astronomical catalogues & associated data:

- Reference catalogues & surveys of astronomical sources at all wavelengths (e.g. SDSS, 2MASS, UCAC, WISE)
  - Tables from papers published in the major astronomical journals
  - A large variety of astronomical data: astronomical sources , spectra, polarization data, but also models (evolutionary, populations, synthetic spectra, ...), statistical analyses, compilations, ...
- ... but with homogenized descriptions (metadata)

The metadata describes the data:

- Type of data (catalog, spectra, light curves etc...)
- Units
- Wavelengths / filters / instrument
- Which object / Region of sky

... in a homogeneous way across all catalogs

**Goal: Discoverability, reusability,**  
provide useful services



Examples:

- Full vizier query (query on ALL >14000 catalogs)
- Photometry viewer
- Associated data



**Find catalogs among 13865 available**

Clear  Find...

Expand search

? *Catalog, author's name, word(s) from title, description, etc. e.g.: AGN, Veron, I/239, or bibcodes...*

▶ [Search for catalogs by column descriptions \(UCD\)](#) ?

▶ [Search for catalogs containing additional data](#)

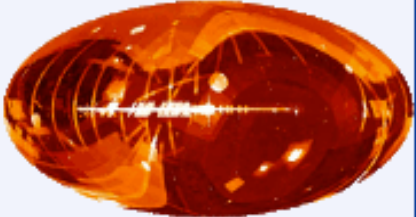
Wavelength	Mission	Astronomy
Radio	AKARI	Abundances
IR	ANS	Ages
optical	ASCA	AGN
UV	BeppoSAX	Associations
EUV	CGRO	Atomic_Data
X-ray	Chandra	Binaries:cataclysmic
Gamma-ray	COBE	Binaries:eclipsing

**Search by Position across 14526 tables**

Target Name (resolved by [Sesame](#)) or Position: Clear  J2000

Target dimension: 2 arcsec Go!

Radius  Box size



[More about VizieR](#) ≤ 375 matching catalogs Find Catalogs ↻

Examples:

- Full vizier query (query on ALL >14000 catalogs)
- Photometry viewer
- Associated data



<input type="checkbox"/>	Reset All	Show table details	or	Query selected Catalogs		
ALL						
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">V/139</a>	(density 70) 794M	(c) The SDSS Photometric Catalog, Release 9 (Adelman-McCarthy+, 2012)	<a href="#">Similar Catalogs</a>	<a href="#">2012ApJS...203...21A</a> <a href="#">ReadMe+ftp</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">II/294</a>	(density 58) 65M	(c) The SDSS Photometric Catalog, Release 7 (Adelman-McCarthy+, 2009)	<a href="#">Similar Catalogs</a>	<a href="#">2009ApJS..182..543A</a> <a href="#">ReadMe+ftp</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">II/314</a>	(density 56) 122M	(c) UKIDSS-DR8 LAS, GCS and DXS Surveys (Lawrence+ 2012)	<a href="#">Similar Catalogs</a>	<a href="#">2007MNRAS.379.1599L</a> <a href="#">ReadMe+ftp</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">I/324</a>	(density 55) 1222M	(c) The Initial Gaia Source List (IGSL) (Smart, 2013)	<a href="#">Similar Catalogs</a>	<a href="#">2013yCat.1324....0S</a> <a href="#">ReadMe+ftp</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">II/319</a>	(density 55) 140M	(c) UKIDSS-DR9 LAS, GCS and DXS Surveys (Lawrence+ 2012)	<a href="#">Similar Catalogs</a>	<a href="#">2007MNRAS.379.1599L</a> <a href="#">ReadMe+ftp</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">I/297</a>	(density 53) 1117M	(c) NOMAD Catalog (Zacharias+ 2005)	<a href="#">Similar Catalogs</a>	<a href="#">2004AAS...205.4815Z</a> <a href="#">ReadMe+ftp</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y	(c) The USNO-B1.0 Catalog (Monet+ 2003)			<a href="#">ReadMe+ftp</a>	



## Examples:

- Full vizier query (query on ALL >14000 catalogs)
- **Associated data**
- Photometry viewer



**Find catalogs among 13865 available**

Clear  Find...

Expand search

? *Catalog, author's name, word(s) from title, description, etc. e.g.: AGN, Veron, I/239, or bibcodes...*

▶ **Search for catalogs by column descriptions (UCD)** ?

▼ **Hide catalogs containing additional data**

time serie  spectrum  images  cube  SED (Spectral Energy Distribution)  none

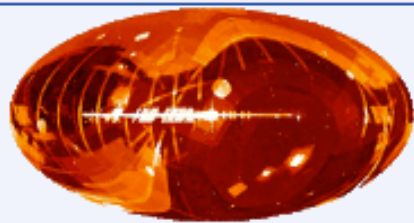
Wavelength	Mission	Astronomy
Radio	AKARI	Abundances
IR	ANS	Ages
optical	ASCA	AGN
UV	BeppoSAX	Associations
EUV	CGRO	Atomic_Data
X-ray	Chandra	Binaries:cataclysmic
Gamma-ray	COBE	Binaries:eclipsing

**Search by Position across 14526 tables**

Target Name (resolved by [Sesame](#)) or Position:

Clear  J2000  arcsec  Go!

Radius  Box size



[More about VizieR](#) ≤ 10 matching catalogs


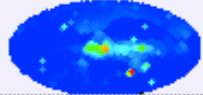

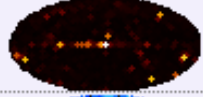
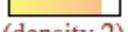
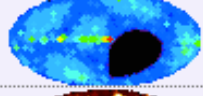

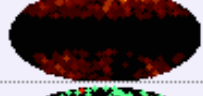

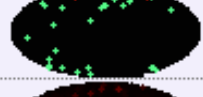
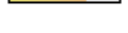

Examples:

- Full vizier query (query on ALL >14000 catalogs)
- Associated data
- Photometry viewer



**Catalog Selection Page**

8 catalogs found having potential matches (6 really found)

<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">B/hst</a>	 (density 11) 466k	(c) HST Archived Exposures Catalog (STScI, 2007)	<a href="#">ReadMe+ftp</a> <a href="#">image/fits</a> <a href="#">Similar Catalogs</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <span style="background-color: orange;">new</span> <a href="#">VI/139</a>	 (density 2) 37k	(c) Herschel Observation Log (Herschel Science Centre, 2013)	<a href="#">ReadMe+ftp</a> <a href="#">image/jpg</a> <a href="#">Similar Catalogs</a> <a href="#">2013yCat.6139...0H</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">VIII/65</a>	 (density 2) 1M	(c) 1.4GHz NRAO VLA Sky Survey (NVSS) (Condon+ 1998)	<a href="#">ReadMe+ftp</a> <a href="#">image/fits</a> <a href="#">Similar Catalogs</a> <a href="#">1998AJ...115.1693C</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">J/A+A/569/A91</a>	 (density 2) 2k	(c) Optical imaging for S4G (Knapen+, 2014)	<a href="#">ReadMe+ftp</a> <a href="#">image/fits</a> <a href="#">Similar Catalogs</a> <a href="#">2014A&amp;A...569A..91K</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <a href="#">VI/112</a>	 (density 2) 114	(c) Optical Imaging of 57 spiral galaxies (Knapen+ 2004)	<a href="#">ReadMe+ftp</a> <a href="#">image/fits</a> <a href="#">Similar Catalogs</a> <a href="#">2003MNRAS.344..527K</a>	
<input type="checkbox"/>	Radmm IR Opt UV X Y <span style="background-color: orange;">new</span> <a href="#">B/chandra</a>	 (density 2) 15k	(c) The Chandra Archive Log (CXC, 1999-2014)	<a href="#">ReadMe+ftp</a> <a href="#">image/fits</a> <a href="#">Similar Catalogs</a> <a href="#">2002yCat...102007C</a>	
<input type="checkbox"/>	<a href="#">Reset All</a>		<a href="#">Show table details</a> <i>or</i> <a href="#">Query selected Catalogs</a>		
<b>ALL</b>					




Examples:

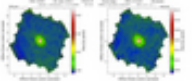
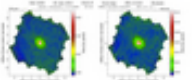

- Full vizier query (query on ALL >14000 catalogs)
- **Associated data**
- Photometry viewer



[VI/139/herschel](#) [Herschel Observation Log \(Herschel Science Centre, 2013\)](#) [2013yCat.6139...0H](#) [ReadMe+ftp](#) 

[Post annotation](#) The Herschel Observation Log (37093 rows)

 start AladinLite

<u>Full</u>	<u>r</u> arcsec	<u>OD</u>	<u>Target</u>	<u>RAJ2000</u> "h:m:s"	<u>DEJ2000</u> "d:m:s"	<u>Proposal</u>	<u>Dur</u> s	<u>Start</u> s	<u>PACS</u>	<u>HIFI</u>	<u>SPIRE</u>
<u>1</u>	0.14	1329	NGC 4030	12 00 23.64	-01 06 00.0	<a href="#">OT1 lcortese 1</a>	397	2013-01-02T00:21:52			
<u>2</u>	0.14	1329	NGC 4030	12 00 23.64	-01 06 00.0	<a href="#">OT1 lcortese 1</a>	397	2013-01-02T00:14:09			
<u>3</u>	0.14	232	NGC 4030	12 00 23.64	-01 06 00.0	<a href="#">KPGT seales01 1</a>	1016	2010-01-01T08:47:05			

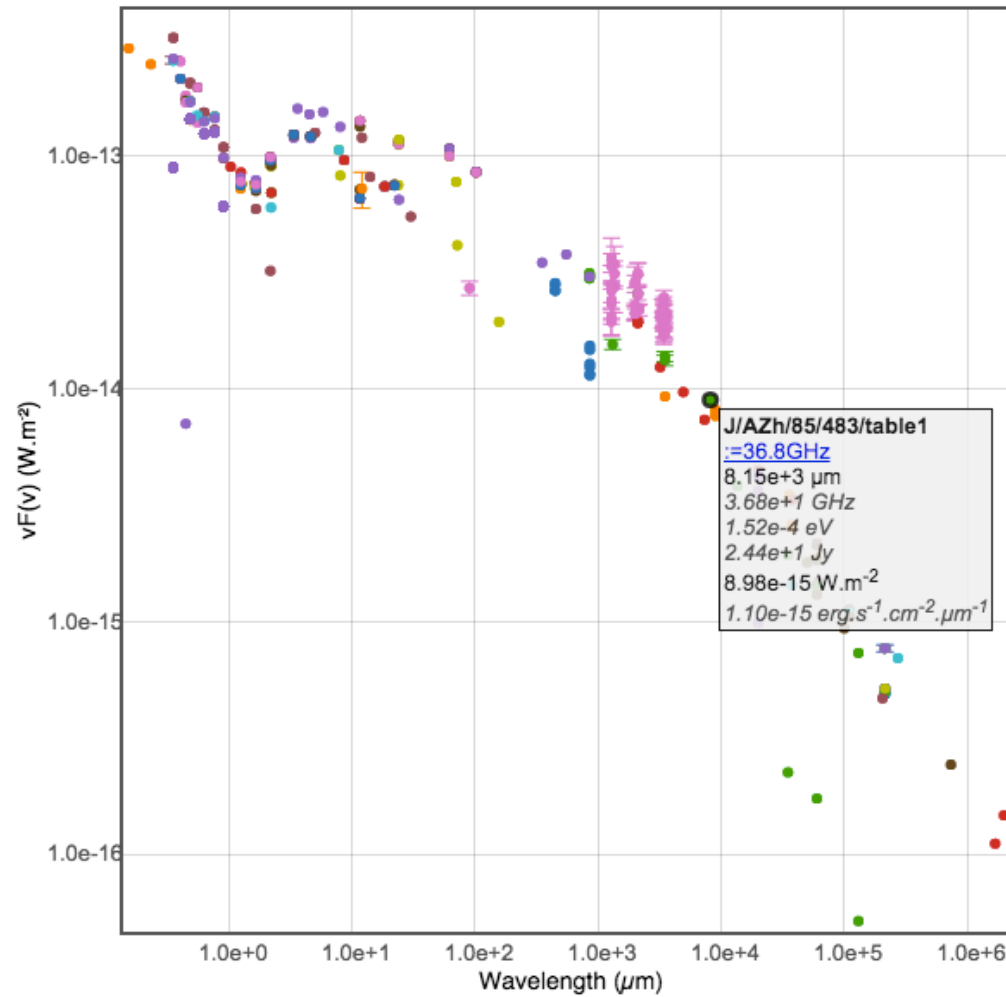
[settings](#) [shar](#)

Target   
 Radius (in arcsec)

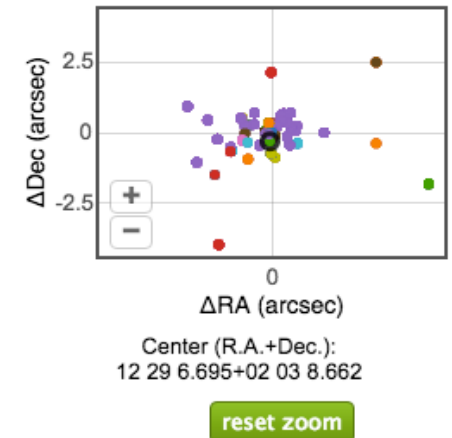


3C273 (12 29 6.695+02 03 8.662),  
radius : 5 arcsec

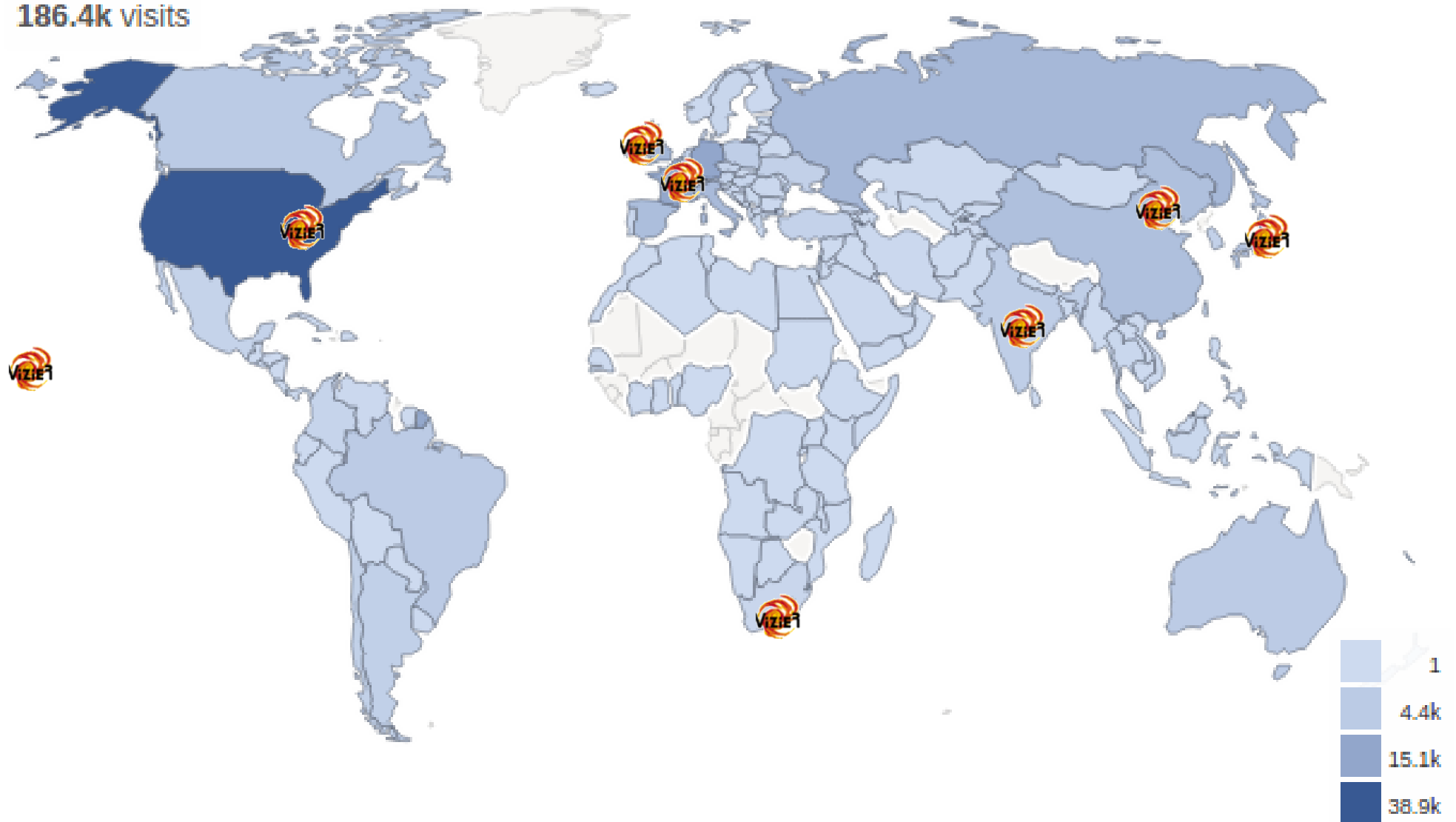
## Photometry viewer



Mouse position:  
 Wavelength :  
 1.93e+1  $\mu\text{m}$   
 Frequency :  
 1.56e+4 GHz  
 Energy :  
 6.44e-2 eV  
 Flux density or  $F(\nu)$  :  
 2.72e-4 Jy  
 $\nu F(\nu)$  :  
 4.23e-17  $\text{W.m}^{-2}$   
 $F(\lambda)$  :  
 2.20e-15  $\text{erg.s}^{-1}.\text{cm}^{-2}.\mu\text{m}^{-1}$

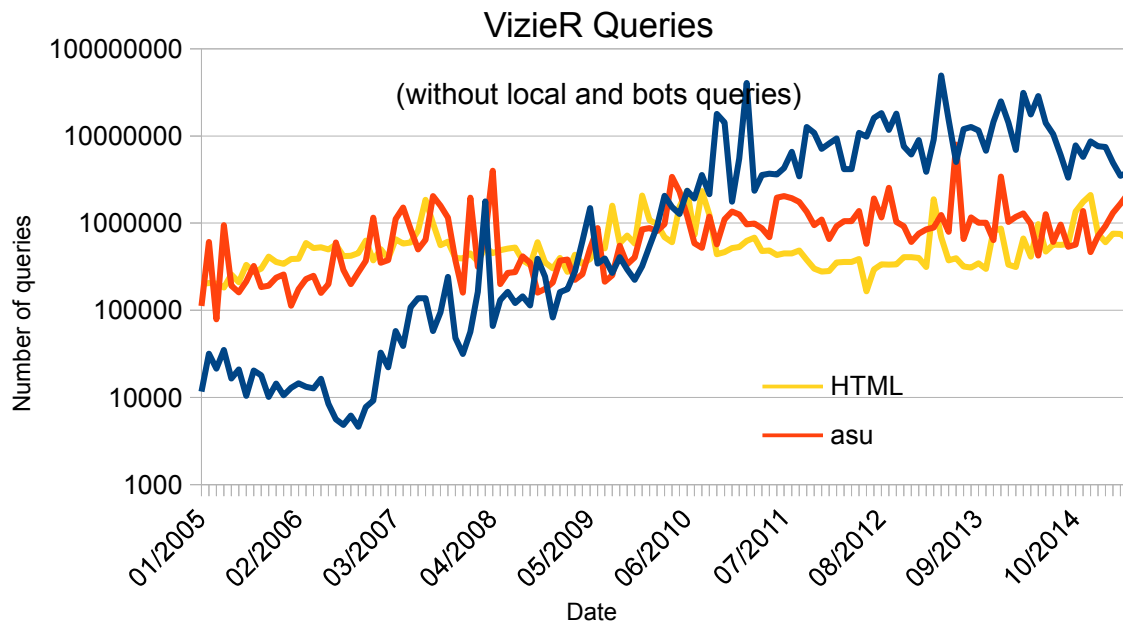
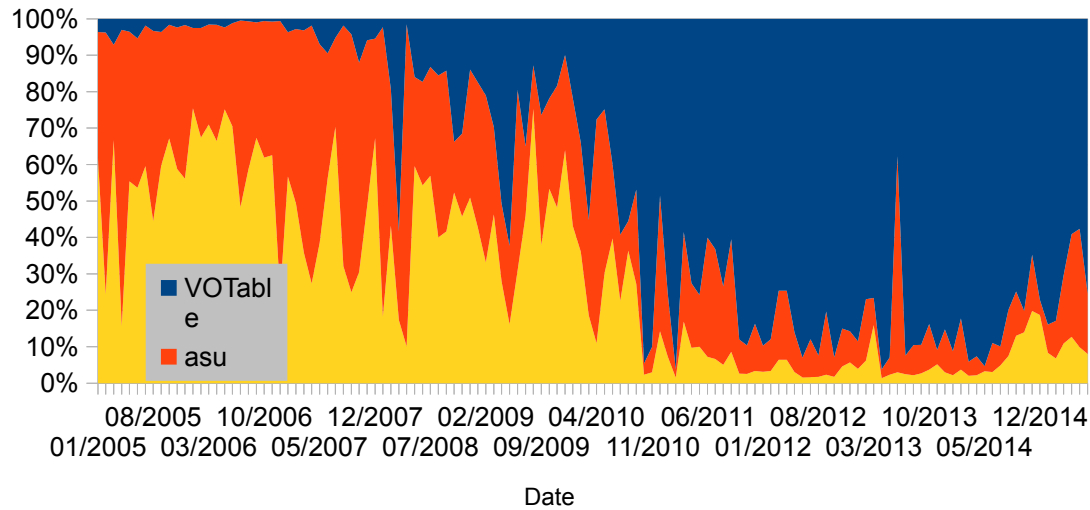


186.4k visits



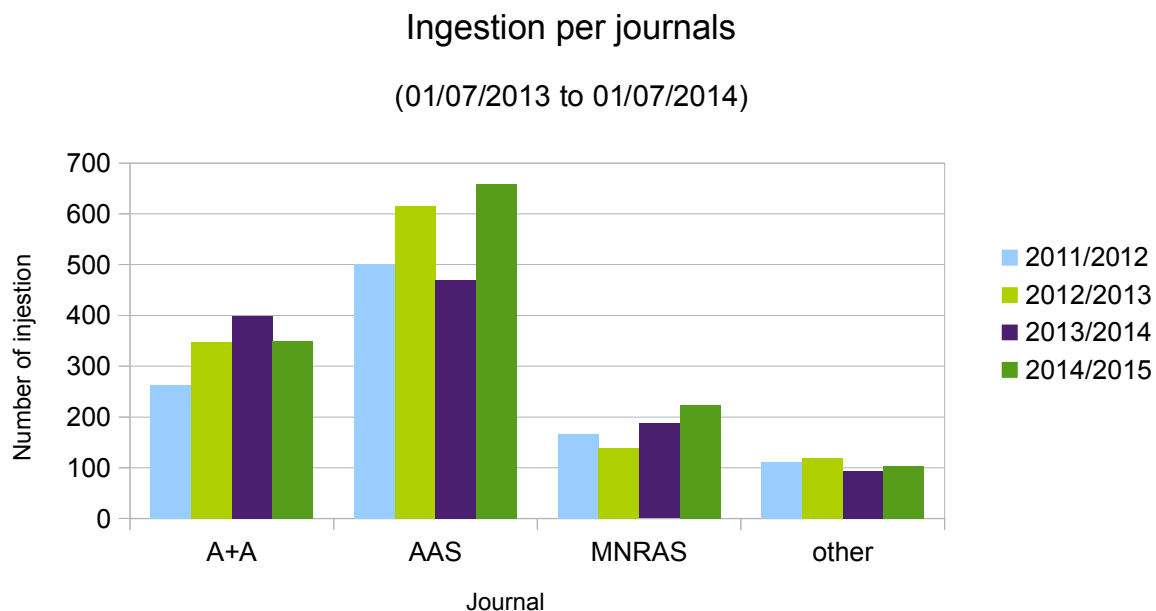


### VizieR output type repartitions



~300,000  
queries/day

## Volume



## Heterogeneity/complexity

 Portal Simbad VizieR Aladin X-Match Other Help

### CDS Catalogues with additional Material

1. [Catalogues with cube \(88\)](#)
2. [Catalogues with filter \(17\)](#)
3. [Catalogues with image \(212\)](#)
4. [Catalogues with model \(88\)](#)
5. [Catalogues with profile \(42\)](#)
6. [Catalogues with spectrum \(543\)](#)
7. [Catalogues with timeSerie \(1451\)](#)



# The paper legacy





# The paper legacy

Page size constraint => squeezed data

- No coordinates
- Truncated/wrong identifiers
- Heterogeneous tabular data
- ...



## Specifications concerning designations for astronomical radiation sources outside the solar system

*A look at the current literature reveals that unclear, ambiguous or confusing designations of astronomical sources of radiation are too often encountered. Therefore, all contributors to databases, and authors of papers, catalogs and surveys, are urged to adhere to the following set of specifications (developed and endorsed by the International Astronomical Union); otherwise, significant data may be irretrievably lost.*

### 1. General recommendations

All source listings should always contain positional information and/or a second designation next to a principal designation in order to avoid ambiguities that can arise with a single designation.

IAU Commission 5 recommendations (November 2006)  
<http://cdsweb.u-strasbg.fr/Dic/iau-spec.html>

# The paper legacy

- No coordinates
- **Truncated/wrong identifiers**
- Heterogeneous tabular data



# The paper legacy



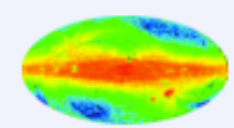
- No coordinates
- Truncated/wrong identifiers


**Table 5**  
Summary of Obscured AGN Candidates Observed with Keck/DEIMOS

Name	R.A.	Decl.	<i>I</i> (mag)	W2 (mag)	W1–W2	$\hat{a}$	$E(B - V)$	<i>z</i>	Q
W1427+3400	14:27:54.57	34:00:43.31	21.90	15.81	0.82	$0.386 \pm 0.091$	$0.41 \pm 0.11$	1.293	A
W1427+3403	14:27:47.16	34:03:41.84	21.24	15.76	0.50	$0.209 \pm 0.117$	$0.33 \pm 0.12$	1.137	A
W1427+3408	14:27:17.93	34:08:28.60	21.71	14.99	2.08	$0.989 \pm 0.009$	$0.88 \pm 0.04$	1.158	A

[II/328/allwise](#) [AllWISE Data Release \(Cutri+ 2013\)](#) [2014yCat.2328....0C](#) [ReadMe+ftp](#)

[Post annotation](#) The AllWISE data Release; please [acknowledge](#) the usage of the WISE data products (747634026 rows) [\(Note\)](#)



 [start AladinLite](#)

<i>Full</i>	<i>r</i>	<i>RAJ2000</i>	<i>DEJ2000</i>	<i>AllWISE</i>	<i>RAJ2000</i>	<i>DEJ2000</i>	<i>Im</i>	<i>W1mag</i>	<i>e</i>	<i>W2mag</i>	<i>e</i>	<i>V</i>
	arcsec	"h:m:s"	"d:m:s"		deg	deg		mag	mag	mag	mag	
<i>1</i>	0.225	14 27 54.565	+34 00 43.09	J142754.56+340043.0	16.9773554	+34.0119702	<i>Im</i>	16.511	0.064	15.714	0.095	

# The paper legacy



- No coordinates
- Truncated/wrong identifiers

## 1. General recommendations

All source listings should always contain positional information and/or a second designation next to a principal designation in order to avoid ambiguities that can arise with a single designation.

## 2. Case of existing designations

- When existing designations are used in listings, they **should never be altered** (e.g., neither truncated, nor rounded, nor shortened).
- In a publication, the bibliographical reference for the designation should be given.

IAU Commission 5 recommendations (November 2006)  
<http://cdsweb.u-strasbg.fr/Dic/iau-spec.html>

# The paper legacy

- No coordinates
- Truncated/wrong identifiers



CDS
Portal
Simbad
VizieR
Aladin
X-Match
Other ▾
Help

## Dictionary of Nomenclature of Celestial Objects

(Last update: 30-Jul-2015)

Designations of astronomical objects are often confusing. Astronomical designations (also called Object *Identifiers*) have been collected and published by Lortet and collaborators in *Dictionaries of Nomenclature of Celestial Objects outside the solar system* ([Biblio](#)). This *Info service* is the electronic look-up version of the *Dictionary* which is updated on a regular basis; it provides full references and usages about 22343 different acronyms.

To find out the meaning of specific acronyms or related references, choose and fill the form below; the words you type in the box are *anded*, i.e. the acronyms matching *all words* will be displayed.

---

Type in your words:

Choose one of the *Info* possibilities:

Select the output layout:

Default
  SIMBAD Usage

Select the number of answers:

...and ask for the result



# The paper legacy

- No coordinates
- Truncated/wrong identifiers



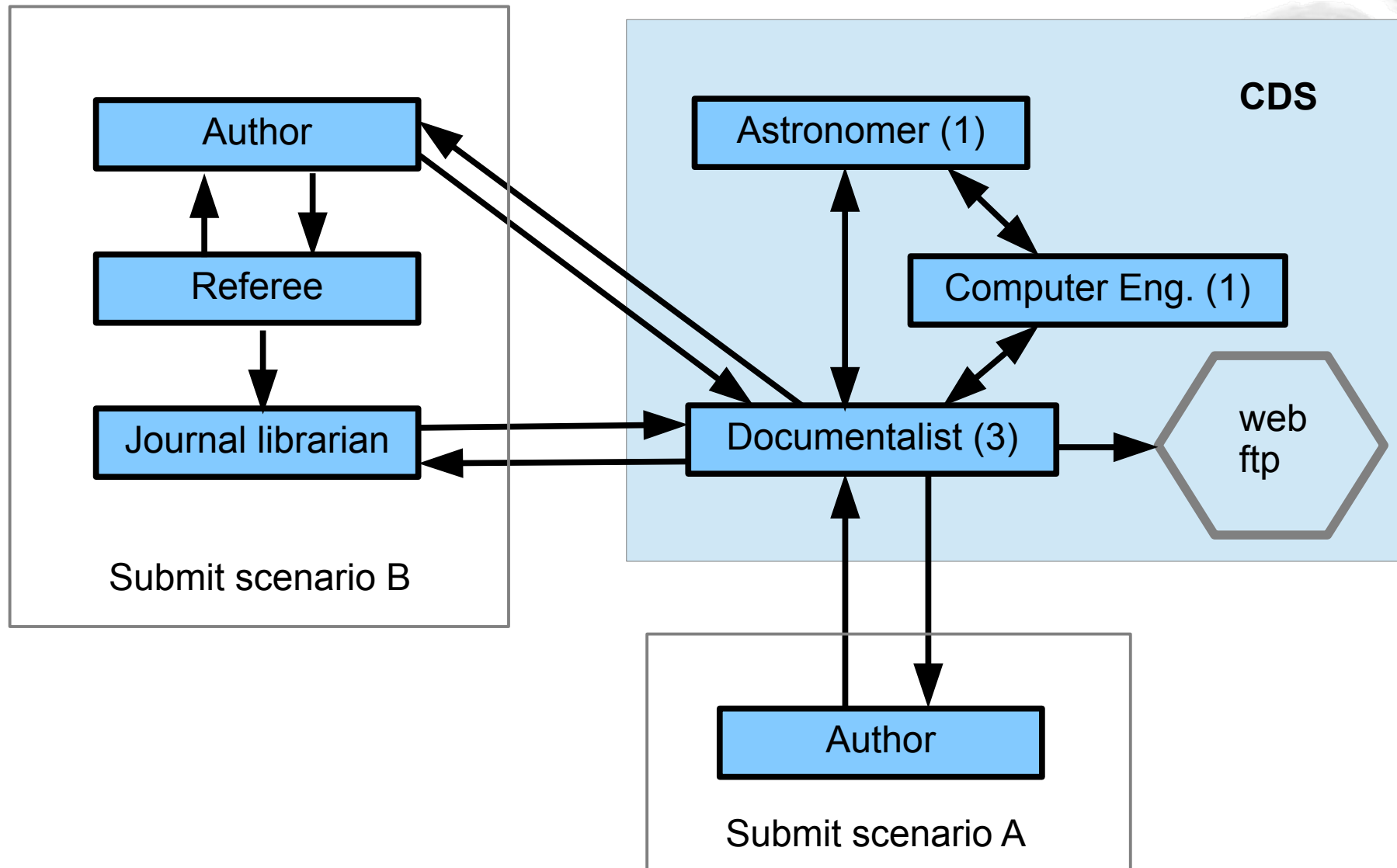
Portal Simbad VizieR Aladin X-Match Other ▾ Help

## Dictionary of Nomenclature of Celestial Objects (Last update: 30-Jul-2015)

Result of query: info cati WISE

Acronym	Use <a href="#">Format</a>	Year	1st Author	<a href="#">Obj. Type</a>
<a href="#">WISE</a>	WISE JHHMMSS.ss+DDMMSS.s	2012	CUTRI R.M.+	(IR)
<a href="#">WISEA</a>	WISEA JHHMMSS.SS+DDMMSS.S	-	Roc Cutri	(IR)
<a href="#">WISEAR</a>	WISEAR JHHMMSS.ss+DDMMSS.s	-	Roc Cutri	(IR)
<a href="#">WISEF</a>	WISEF JHHMMSS.SS+DDMMSS.S	-	Roc Cutri	(IR)
<a href="#">WISENF</a>	WISENF JHHMMSS.ss+DDMMSS.s	2010	WRIGHT E.L.+	(IR)
<a href="#">WISEP</a>	WISEP JHHMMSS.ss+DDMMSS.s	-	Roc Cutri	(IR)
<a href="#">(WISEPC)</a>	WISEP JHHMMSS.ss+DDMMSS.s	-	Roc Cutri	(IR)
<a href="#">WISEPF</a>	WISEPF JHHMMSS.SS+DDMMSS.S	-	Roc Cutri	(IR)
<a href="#">WISER</a>	WISER JHHMMSS.SS+DDMMSS.S	-	Roc Cutri	(IR)

**Note:** Usage of acronyms in parentheses like 'WISEPC' should be *avoided*



## Roles at CDS



Providing access to useful, discoverable, reusable scientific data requires:

- **Astronomer** : scientific expertise and watch
- **Computer eng.** : develop/maintain efficient means for ingestion, retrieval, use of data
- **Documentalists** : generate metadata, added value, consistency checks, interact with authors

## Roles, **new roles?**



- **Author** : produces sci. data, **consistency checks?**
- **Referee** : validates sci. results, **consistency checks?**
- **Journal librarian** : formats and posts tabular data, **consistency checks?**
- **CDS** : provides and maintains tools for consistency checks (e.g. dictionary of nomenclature)

Leaving paper behind: electronic tables **NEED NOT**,  
**SHOULD NOT** be copies of paper tables