

Strasbourg astronomical data centre (CDS)

42 years serving the community

Francoise Genova

An old (a very old) data centre!

- A 42-year old history
- Founded by INAG – now CNRS-INSU (in charge of French ground-based astronomy) in 1972, in agreement with the Université Louis Pasteur – now Université de Strasbourg
- *A far-seeing vision*
 - Collect ‘useful’ data on astronomical objects, in electronic form
 - Improve them by critical evaluation and combination
 - Distribute the results to the international community
 - Conduct research using the data
- Users ARE trusting us (~1 million queries/day)
- But external certification is useful
 - Labelled as “Research Infrastructure” at the national level (as well as telescopes...)
 - Regular member of the World Data System
 - Awarded the Data Seal of Approval



- On-line data and services: observatory archives, value-added services and tools (CDS), electronic journals and the NASA ADS bibliographic database, ...
- On-line resources are widely used by the community
- Major scientific objectives such as multi-wavelength/multi-technique astronomy to understand the physics at work in the objects, study variable/long term phenomena, etc
- Optimizing the scientific return of the ground- and space-based instruments
- Astronomy has been a pioneer of « Open data » (proprietary period when competitive process to get observation time)
- Long term international partnership for defining standards to describe data and network on-line resources
- CDS a pioneer and one of the major actors at the international level

- Support the international astronomical community in its research tasks (not only collect/preserve information), *science driven* (not technically driven)
 - Core task: provide highly used value-added services
 - Keywords: quality, scientific and technical relevance, collaborations, global networking of expertise and resources
 - Staff: an integrated team of scientists, data librarians and computer engineers (~30 staff members, 1/3 on each profile)
- Close collaboration with academic journals, space and ground-based telescope archives, ADS, other data centres
- Support to projects, including a major role in the Virtual Observatory (set of interoperability standards and tools enabling seamless access to on-line data)

- Reference, added-value services
 - The basis of CDS service to the community
- Technological watch, R&D
 - One key of long term sustainability
 - The outcome: new services, new functionalities for the users and for ourselves
- Virtual Observatory
 - Interoperability standards, tools (eg: the VO portal for images)
- Participation in projects
 - E.g. Gaia, RDA
- User support, knowledge dissemination, expertise
- Science
 - Scientific staff are ACTIVE scientists
 - Another key of long term sustainability



Astronomical objects

identification, bibliography,
data, measurements



Federation of tabular data

catalogues, published tables
observation logs, very large surveys



Interactive sky atlas: Data discovery, integration, visualisation, manipulation

images, databases, catalogues,
surveys, archives, *user data*

Among the challenges

SIMBAD	2008	2011	2014
Objects	3.900.000	5.400.000	7.750.000
Identifiers	11.750.000	15.124.000	18.551.617
References	216.000	253.000	294.000
Citations	5.750.000	8.180.000	10.740.000

VizieR	2008	2011	2014
Catalogues	6.500	9.100	12.650

- Growth in volume and complexity
- Rapid evolution of astronomy & technology
- Operational constraints (24/24, 7/7, significant usage increase)
- Long term sustainability
- Agile strategy needed, taking all aspects into account

- Circ. 1994, we needed to use bibliographic references as a link
- A disciplinary bibcode which had been defined before the internet era between CDS and a US database, because we exchanged information, then also used by the ADS [2014A&A...561A.108B](#)
- Also used by the on-line journals to network bibliographic information ... now DOI...
- SIMBAD client – server
- Allows a distant software to query SIMBAD and reuse the result, eg to translate an object name into a position or a list of references
- Development decided in 1992
- A bold decision since the CDS is hidden under the up-front service
- But it was worth!

SIMBAD usage in ADS and telescope archives



ADS Labs SIMBAD object types from a list of references

Simbad interface



SIMBAD Name Resolver usage in ESO Archive

If you would like to query the Archive for instrument specific parameters, please use the [dedic](#) to search for **reduced Data Products**, please have a look at the [ESO Data Products](#) page and [the checkboxes on the right of the parameters define whether or not they will be displayed on](#)

SEARCH ShowAll ShowNone Reset

Target, Program and

Target Name Resolved by SIMBAD

RA DEC J2000

Search Box Input RA(h) DEC(deg)

Output Sexagesimal (h, deg)

List of Targets Parcourir...

in ESA ISOC Science Data Archive

Search MAST for a Target or Mission

Enter Target name (or Coordinates):

Resolver: SIMBAD NED Don't Resolve

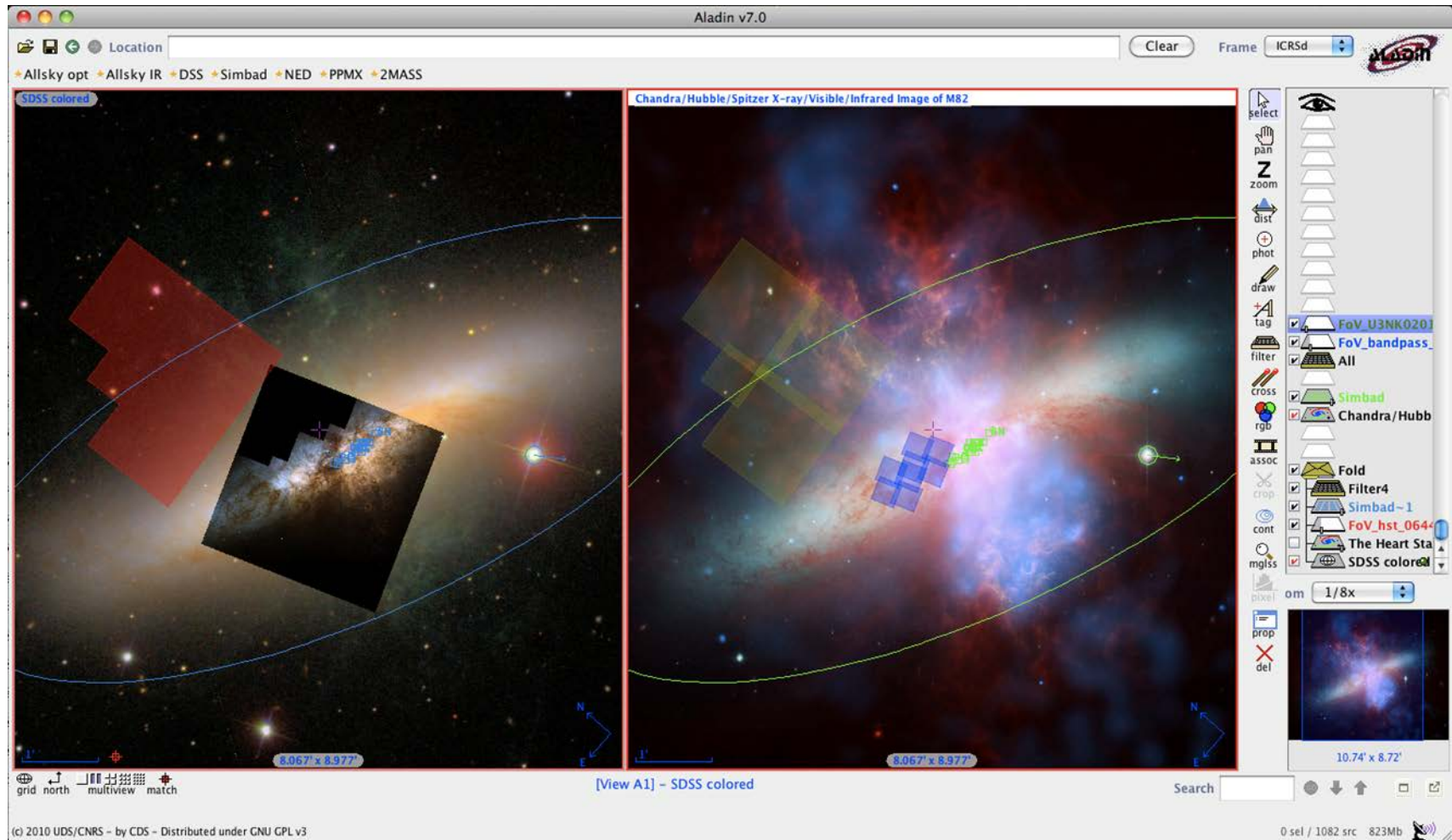
and/or Band/Data Type(s): [more options](#)

	Extreme UV	Far UV	Near UV	Optical	Near IR	Radio
Images	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Spectra	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Search Reset Help

in STScI MAST Archive

Integration of observational data



Images are built on-the-fly with data extracted from remote observatory archives

VO allows interoperability of tools

Interoperability

Aladin

Topcat

Star	RA	Dec	Mag	Class	Dist
31	1128+704	MR 280	8.044	14.28	
32	1146+827	PKS	8.341	26.9	
33	1146+549	PC	6.969	18.82	
34	1156+291	4C 28.65	6.729	14.41	
35	1160+281	PC	6.185	15.31	
36	1212+143	PC	6.093	14.85	
37	1218+763	MR 285	6.07	14.5	
38	1225+317	82	6.319	16.67	
39	1226+254	IC 263	6.516	16.28	
40	1228+204	T06 1342	6.044	13.7	
43	1242+170	PC	7.271	18.38	
45	1253+658	4C 2279	6.559	17.75	
48	1167+382	PKS	6.223	14.52	

VOSpec

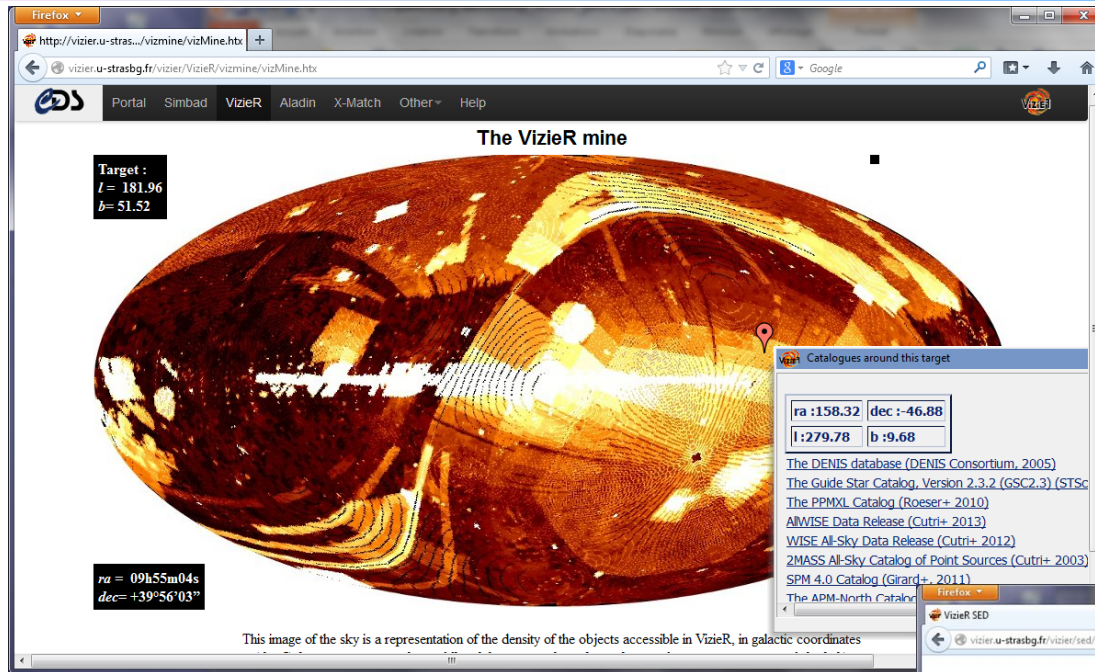
Your programs

+

...

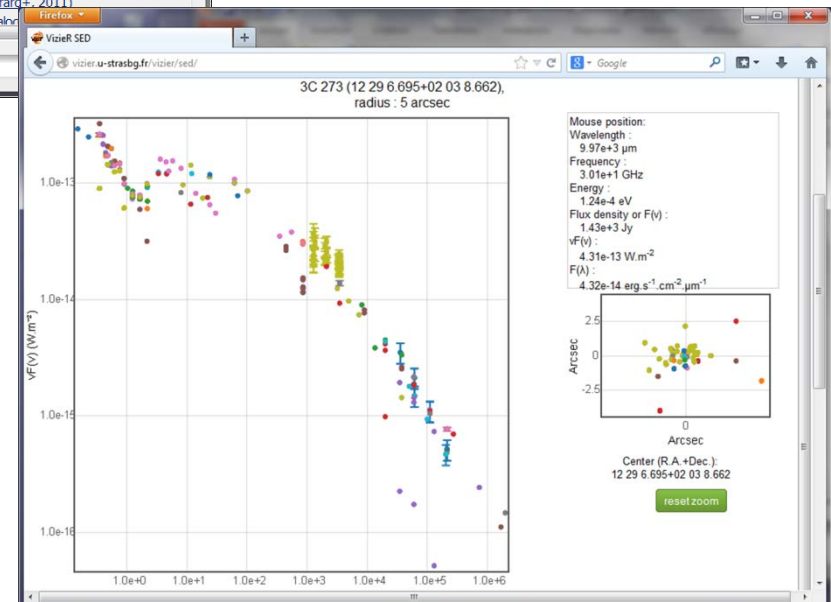
But also long tail, at least a kind of...

- CDS hosts data « attached » to academic publications: tables, images, etc, in VizieR which also hosts object catalogues from ground- and space based observatories (up to 2 billion rows)
- Agreement between CDS and the academic journals (started in 1993 with *Astronomy & Astrophysics*)
- Starting point: tabular data from publications
 - **Printed figures became usable data!**
- More and more non-tabular data « attached to publications » (images, spectra, time series) - A&A: the scientific editor asks the authors to deposit their data
- More than 12 000 « catalogues »
- Standard description, agreed with the journals
- Hundreds of different quantities, well qualified by a UCD (VO standard)
- Currently a persistent ID, *DataCite advice will be sought (link with article ID, granularity,...)*

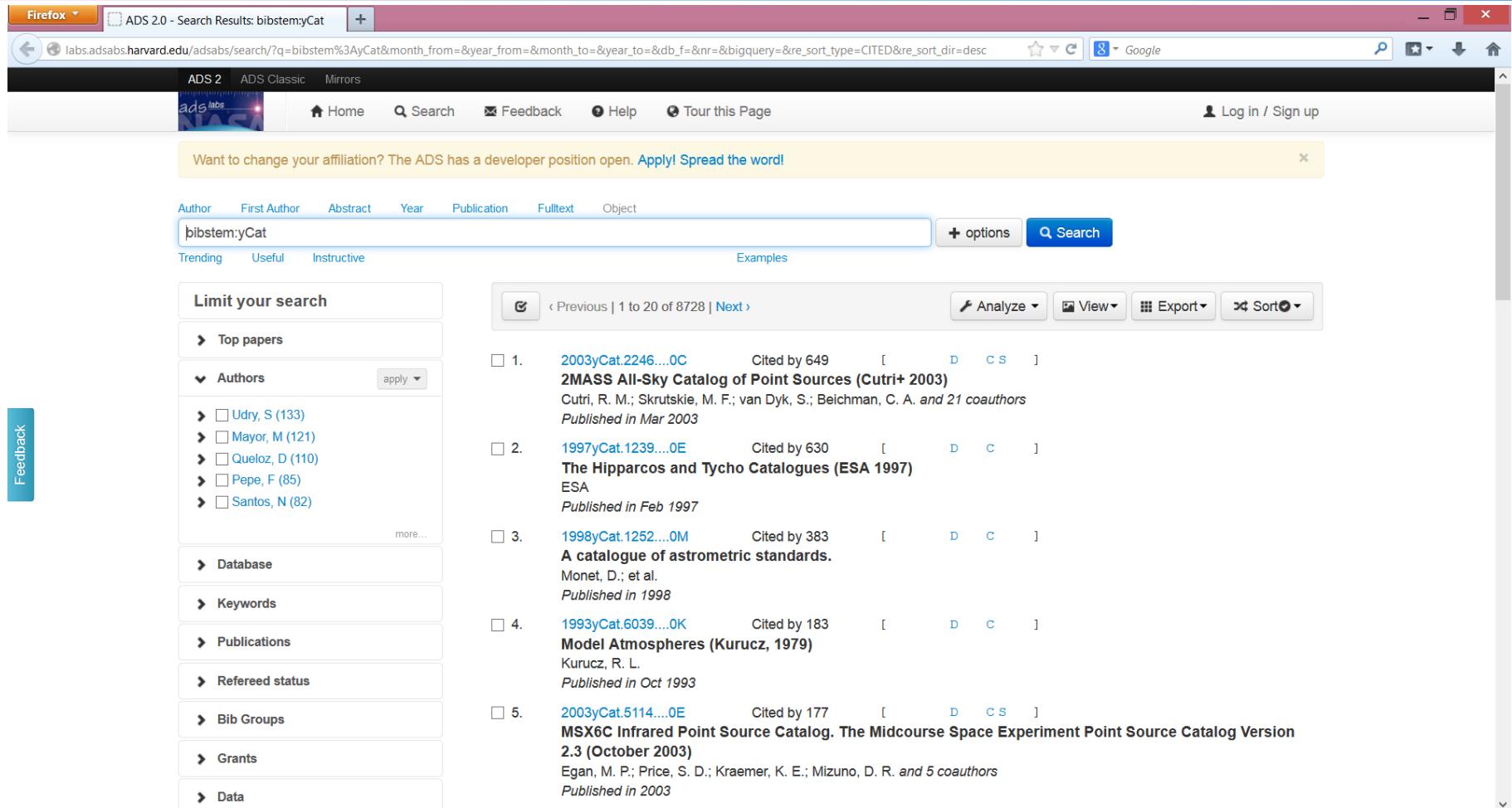


*Data validated
by a publication -
Fully discoverable
(OAI-PMH VO registry)
and usable
including numbers
inside the tables*

“Photometry viewer”:
Spectral points
extracted from the
collection



These datasets are used and cited



The screenshot shows the ADS 2.0 search results page for the query 'bibstem:yCat'. The browser is Firefox, and the URL is 'labs.adsabs.harvard.edu/adsabs/search?q=bibstem%3AyCat&month_from=&year_from=&month_to=&year_to=&db_f=&n_r=&bigquery=&re_sort_type=CITED&re_sort_dir=desc'. The page features a navigation bar with 'Home', 'Search', 'Feedback', 'Help', and 'Tour this Page'. A search bar contains the query 'bibstem:yCat' and a 'Search' button. Below the search bar, there are filters for 'Trending', 'Useful', and 'Instructive'. A 'Limit your search' sidebar on the left lists categories like 'Top papers', 'Authors' (with a list of authors and their counts), 'Database', 'Keywords', 'Publications', 'Refereed status', 'Bib Groups', 'Grants', and 'Data'. The main content area displays a list of search results, each with a checkbox, a link to the dataset, the number of citations, and a list of authors. The results are:

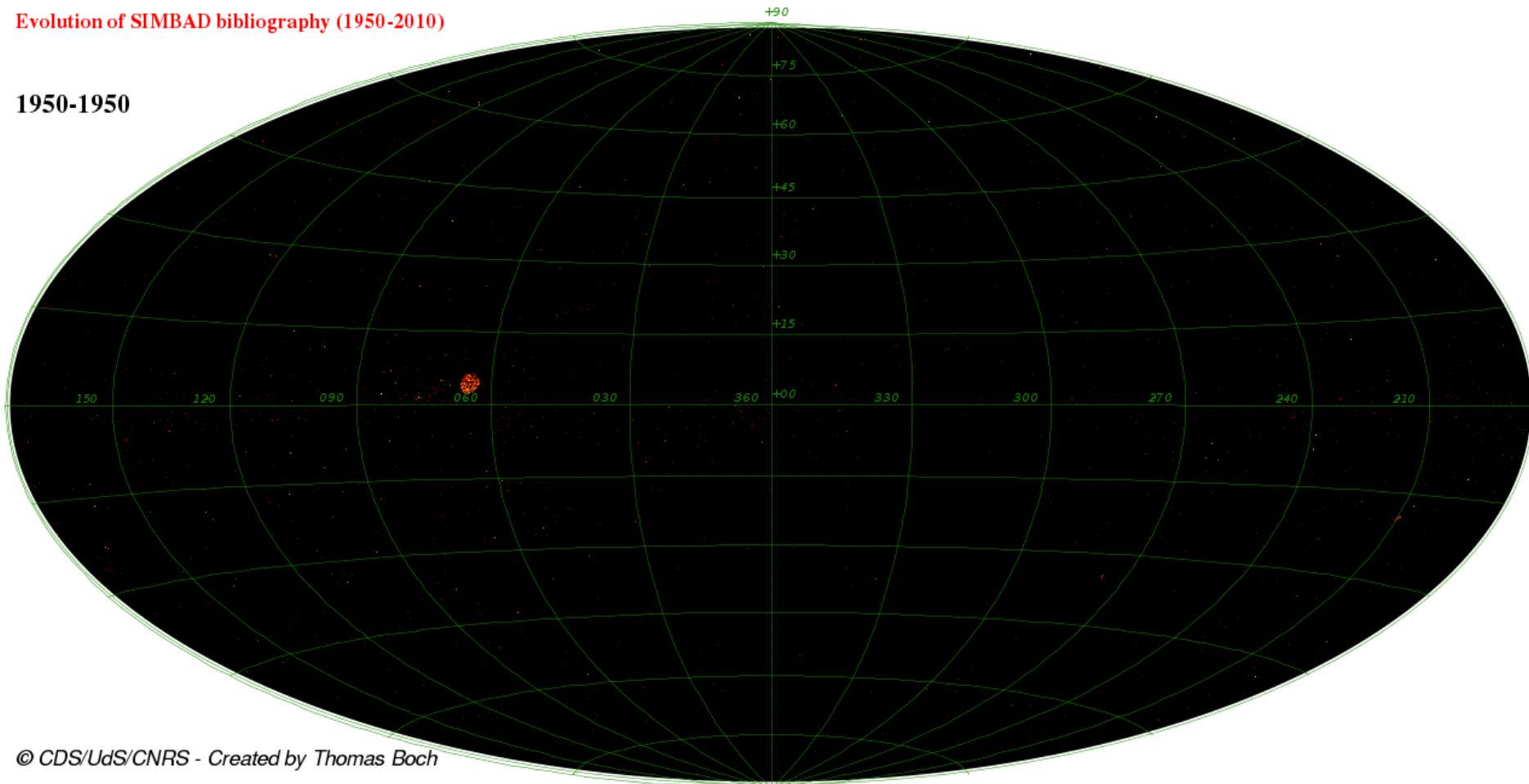
1. [2003yCat.2246....0C](#) Cited by 649 [D C S]
2MASS All-Sky Catalog of Point Sources (Cutri+ 2003)
Cutri, R. M.; Skrutskie, M. F.; van Dyk, S.; Beichman, C. A. and 21 coauthors
Published in Mar 2003
2. [1997yCat.1239....0E](#) Cited by 630 [D C]
The Hipparcos and Tycho Catalogues (ESA 1997)
ESA
Published in Feb 1997
3. [1998yCat.1252....0M](#) Cited by 383 [D C]
A catalogue of astrometric standards.
Monet, D.; et al.
Published in 1998
4. [1993yCat.6039....0K](#) Cited by 183 [D C]
Model Atmospheres (Kurucz, 1979)
Kurucz, R. L.
Published in Oct 1993
5. [2003yCat.5114....0E](#) Cited by 177 [D C S]
MSX6C Infrared Point Source Catalog. The Midcourse Space Experiment Point Source Catalog Version 2.3 (October 2003)
Egan, M. P.; Price, S. D.; Kraemer, K. E.; Mizuno, D. R. and 5 coauthors
Published in 2003

They have an ID in the ADS, and ADS is used by astronomers to prepare the reference list for their publications – not perfect, but a starting point

Building a knowledge base on 40 years

Evolution of SIMBAD bibliography (1950-2010)

1950-1950

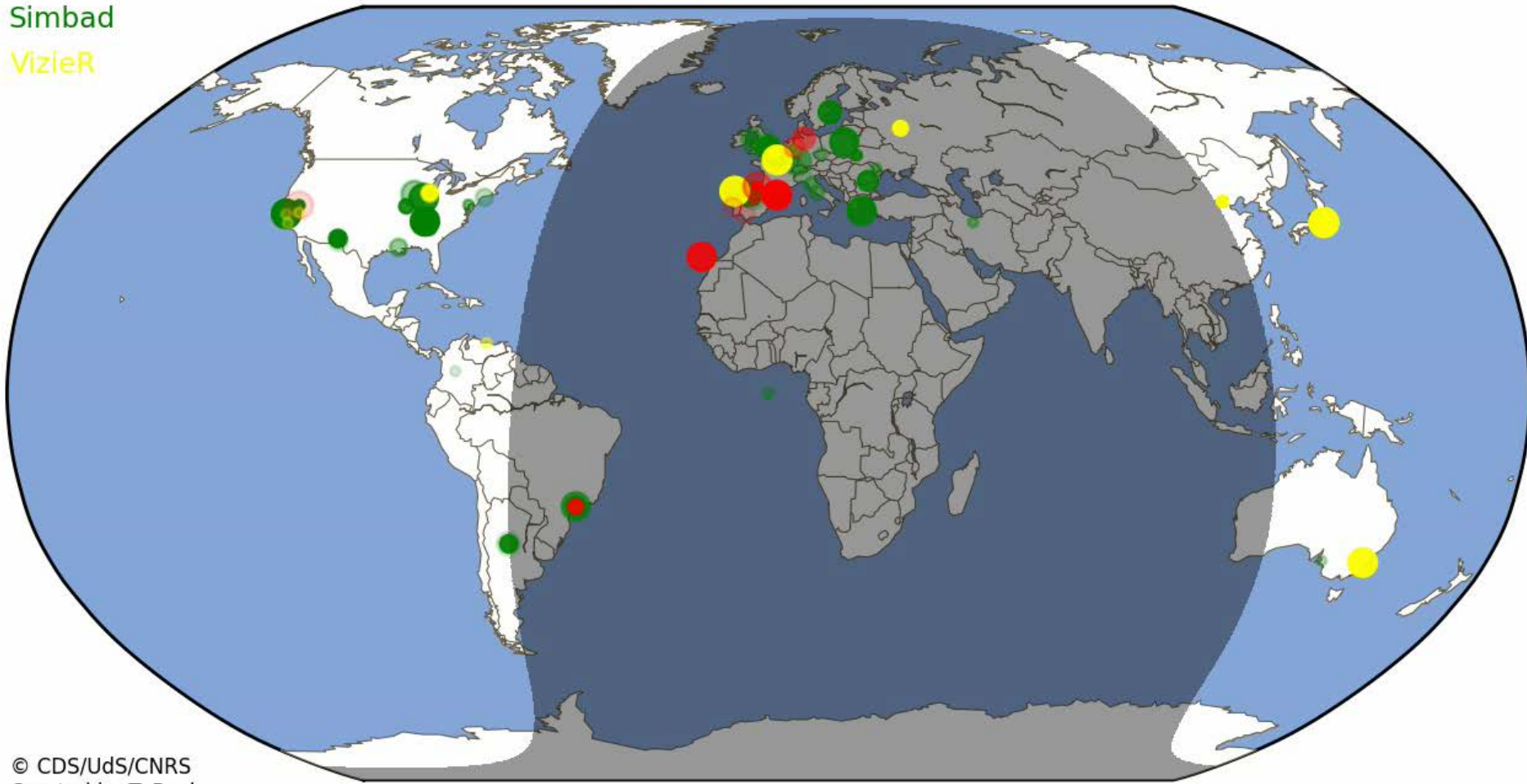


© CDS/UdS/CNRS - Created by Thomas Boch

... and a user base

2012-09-05 22:00:00 UTC

Aladin
Simbad
VizieR



© CDS/UdS/CNRS
Created by T. Boch