



Scientific council 2023

VizieR Staff and contributors:

Astronomers: P.Ocvirk, C. Bot, G. Monari, S.Derriere

Software Engineers: G.Landais, F.X.Pineau,

[A.Flint](#), T.Boch

Documentalists: P.Vannier, E. Perret, [C. Fix](#),

[A.Fiallos](#), M. Brouty



CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

Non-CDS: C. Saillard, T. Keller, L. Michel
(Strasbourg Observatory)

Departures from VizieR



Alicia FLINT

- Ingenieure de Recherche, VizieR developer
- Support for VizieR code renewal
- Left 24/11/2023



Ana FIALLOS

- Ingenieure d'etudes, VizieR documentalist
- Ingestion of catalogues
- Leaving end of Nov. 2023



Coralie FIX

- Ingenieure d'etudes, VizieR documentalist
- Ingestion of catalogues
- Leaving August 2024



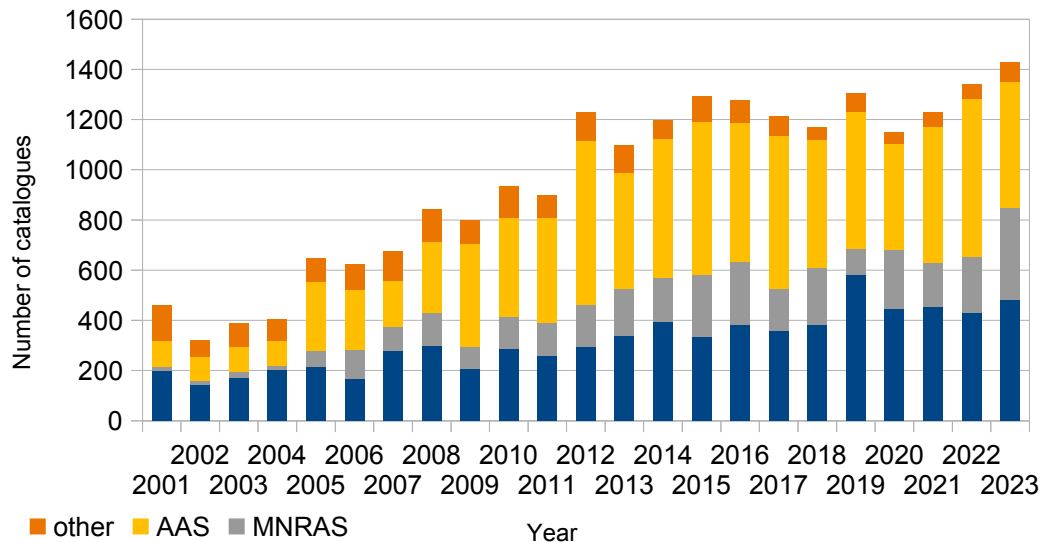
VizieR content - II



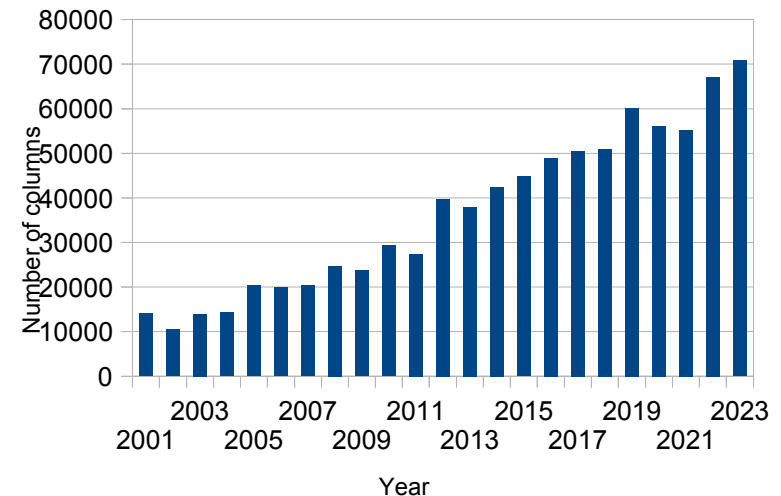
Ingestion statistics :

- A steadily increasing volumetry : +1419 in past year
 - Total (23/11/2023): 24 483 catalogues, >56k tables
- Slower increase of records (post-Gaia DR3) : ~82 billions records

Catalogues evolution

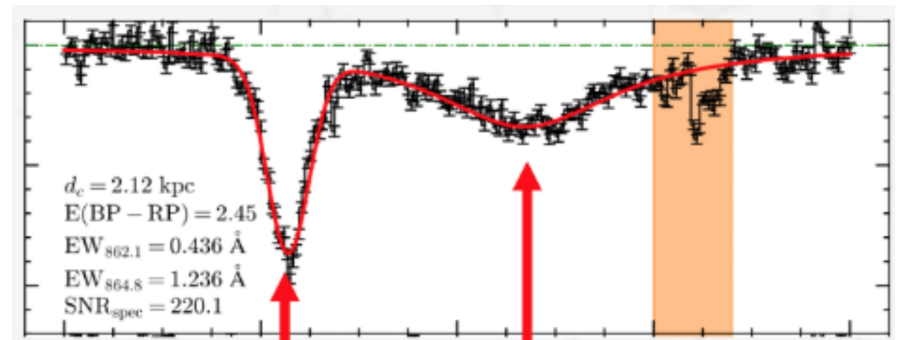
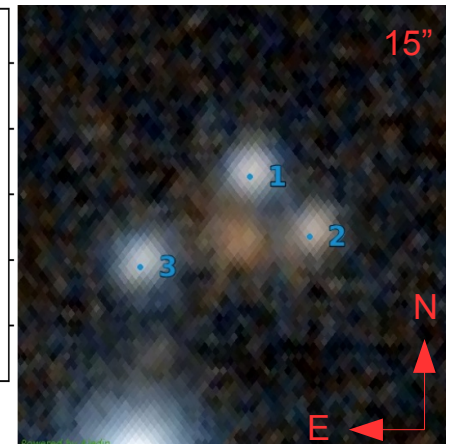
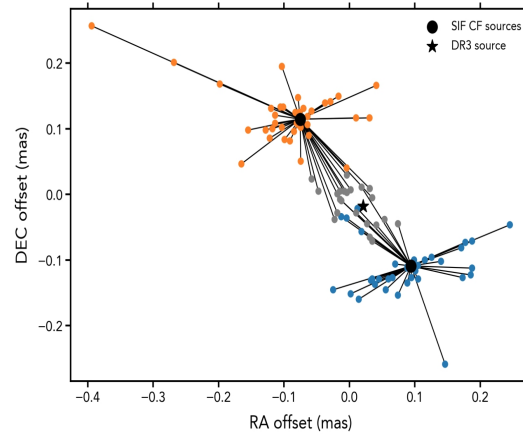


Number of columns evolution



Gaia Focused Products Release

- Data improved thanks to **longer timeline** (66 months instead of 33) and **improved crowded fields** :
 - Omega cen (pos, PM, mags)
 - Lensed quasars
 - Long Period Vars, RVs(t)
 - DIBs stacked spectra
 - Solar system update
- 11 tables
- including 3 large tables
- Up to 171 million records





Very large cats ingested in past year:

- Gaia:

- Carrasco+2023

- Maiz-Appellaniz+2023 (V*)

- Holl+2023

- ESO Phase III:

- VMC DR6

- VVV DR4.2

- Dark Energy Spec. Instrument DR8

“Thick” catalogs: > 150 columns

- Gomez-Munoz 2023 (432 columns):
Pne in GALEX/optical surveys

In progress/standby/contact made:

- ZTF: contact made, standby
- HSC: contact made, standby
- Pan-STARRS DR2: download imminent

Planned for next year:

- DECALS DR9/10
- KIDS DR4
- ESO phase 3: ATLAS-DR4, VPHAS+ DR3.2, VIKING DR4



Accessing VizieR data

VizieR services and access modes: discovery, table access, visualization, photometry, associated data, astroquery, ...

VizieR provides access to the most complete library of published astronomical catalogues and data tables available on line organized in a self-documented database. Query tools allow the user to select relevant data tables and to extract and format records matching given criteria. Currently, 15024 catalogues are available. more info

Free text search: catalogue name, author, ... Find catalogues

Position: position or object name 10 Find catalogues Photometry

Go to the classic form Advanced search

VizieR

- How to publish my catalog
- Help and tutorials
- View large catalogues
- Rules of usage
- Mirrors

Other related services

- TAPVizieR
- Photometry viewer
- CDS cross-match service
- VizieR images, spectra service
- VizieR using the batch mode

Simple browsing modes

- By Hierarchical organisation
- By acronyms or abbreviations
- By popularity
- Recently entered into VizieR
- Catalogues having images, spectra...

VizieR

Search Criteria

Keywords: RA12000 Dec2000 RA12000 Dec2000

Tables: The Pan-STARRS release 1 (PS1) Survey - DR1 (Chambers - 2016) Pan-STARRS DR1 catalogue (original column names in green) (1919300000) Headline by

Columns: RA12000 Dec2000 RA12000 Dec2000 objid

Search

Tables

TAP

Search associated data using the TAP interface

Search criteria

Search

Associated data

VizieR Photometry viewer

Photometry

Plot

Derived products provided in CDS or data available through software, API..



CDS X-Match Service

Choose tables to cross-match

VizieR SIMBAD SIMBAD ZMASC

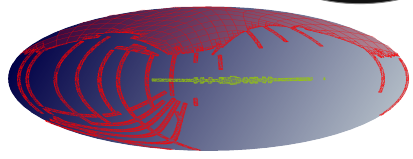
Begin the X-Match

xmatch

cassis

Aladin

topcat

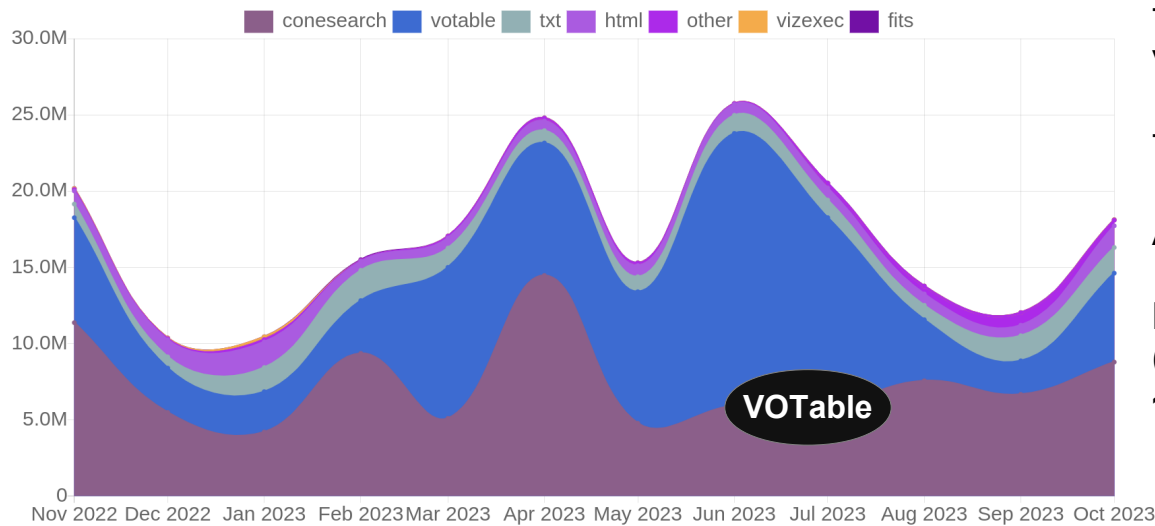


VizieR usage statistics



VizieR statistics (from the CDS statistics collector)

VizieR queries repartition by output (classic web form and HTTP API)



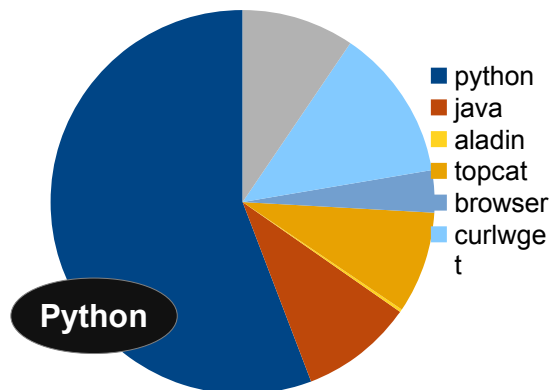
Total : ~750K queries/day
Very strong increase!

TAP : ~69.4K queries/day

Associated data (spectra/images):
~350 queries/day

**Importance of the Virtual Observatory
(conesearch+VOTable)
~86% of the queries**

Queries repartition per user-agent

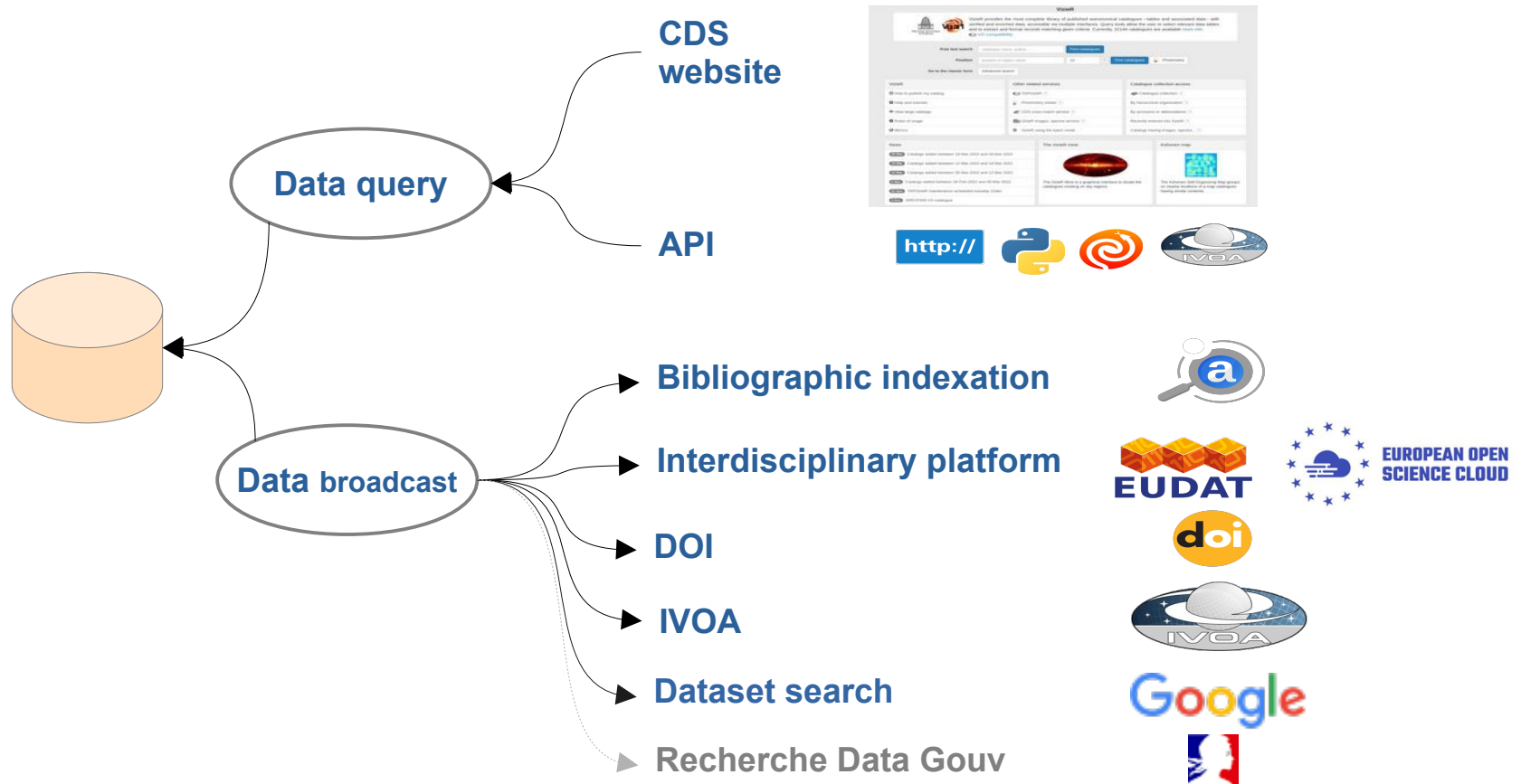


Importance of **Python(56%)**,
TOPcat (9-18%), curl+wget (17%)

VizieR dissemination in IVOA, Open Science



Broadcast catalogues in the community and in **Open Science** (interdisciplinary) networks
→ DOI as a unique and persistent identifier



Is VizieR F.A.I.R. ?



YES from the astronomical community
Not quite true, from Open Science point of view

Enter resource identifier (URL/DOI)

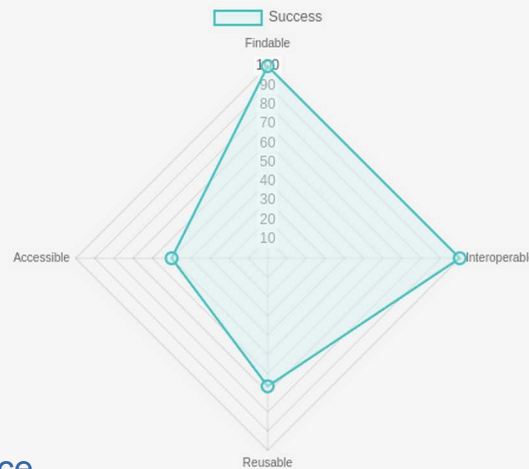
<https://cdsarc.cds.unistra.fr/viz-bin/cat/J/AJ/137/226> ✓ [Test all metrics](#)

The URL/DOI is valid

[Clean results](#)

[Dataset Dataverse](#) [Workflow](#) [Publication Datacite](#) [Dataset](#) [Tool](#)

Radar chart of metrics completion



A1.2 Authorisation procedure or access rights

R1.1 Metadata includes license

<https://fair-checker.france-bioinformatique.fr/check>



VizieR renewal large catalogue access



QATSS (FX.Pineau)

Large tables (>50millions records) are stored in binary file (efficient for conesearch) and in database (TAP access)

- A new architecture to provide Large tables in binary (RCF) format improve the robustness

An original way to manage huge astronomical table

The Foreign Data Wrapper technology

PostgreSQL Foreign Data Wrapper is an implementation of the SQL standard SQL-FDW (Management of External Data).

It explores a cost-effective solution in terms of funds and maintenance which consists in importing ATSS binary data to foreign tables the development of a PostgreSQL extension using the Foreign Data Wrapper that maps SQL with API of the binary file including the cone search and conequery capabilities. The architecture has been tested with a selection of large tables in a TAP service, we will discuss the technology in a poster at the beginning of the current development.

The PostgreSQL FDSW API is available in Python, in Ruby and in C.

R&D work exploiting ATSS robustness

The FDW API

The implementation of a new FDW involves implementing a collection of handlers functions registered in a PostgreSQL "Translation" structure. The handler functions are executed by PostgreSQL at different stages of the database SQL parsing SQL execution, etc.

The API allows PostgreSQL to interact with the Query parser by implementing additional capabilities. The architecture has been tested with a selection of large tables in a TAP service, a constraint on identifiers, or any capabilities offered by the service API.

Conclusion

The coupling of Foreign Data Wrapper and ATSS has been tested. It is an interesting storage alternative adapted to large tables.

This FDW extension is implemented in PostgreSQL. The code in particular for SQL translation has not been developed. It shows a solution offering a SQL API right interface that can be a candidate for handling large datasets in a simple architecture.

G.Landini, F.X.Pineau, P.Ouvry, A.Vizier-Fantaneli, A.Lepetit
glandini@obs.uvst.fr | vizier@obs.uvst.fr | ouvry@obs.uvst.fr | lep@obs.uvst.fr

The Astronomical Table Serialisation System Query tool (QATSS) and its ecosystem

QATSS

QATSS is a query tool that reads query tables in a binary format, the ATSS, and returns the results in a standard format. The architecture is based on the PostgreSQL Foreign Data Wrapper (FDW) technology which allows to import ATSS binary data into PostgreSQL. The architecture has been tested with a selection of large tables in a TAP service, we will discuss the technology in a poster at the beginning of the current development.

ATSS

ATSS is a query tool that reads query tables in a binary format, the ATSS, and returns the results in a standard format. The architecture is based on the PostgreSQL Foreign Data Wrapper (FDW) technology which allows to import ATSS binary data into PostgreSQL. The architecture has been tested with a selection of large tables in a TAP service, we will discuss the technology in a poster at the beginning of the current development.

QATSS features examples

QATSS is a query tool that reads query tables in a binary format, the ATSS, and returns the results in a standard format. The architecture is based on the PostgreSQL Foreign Data Wrapper (FDW) technology which allows to import ATSS binary data into PostgreSQL. The architecture has been tested with a selection of large tables in a TAP service, we will discuss the technology in a poster at the beginning of the current development.

Poster related to ATSS in ADASS (2023)

VizieR renewal

Assign machine-readable column



UCD (Unified Content Descriptor) migration UCD (simple) UCD1+ (complex)

UCD	Score	Origin
phot.flux.density;phys.flux;em.IR.60-100um	9.6	compute
phot.flux.density;em.IR.60-100um	8.6	existing
phot.flux.density	6.1	existing
phys.density	3.1	existing
phot.flux.density;phys.density;em.IR.60-100um	9.6	compute
phot.flux.density;phys.flux;phys.density	8.6	compute

<https://cds.unistra.fr/ucd-finder/beta/>

IVOA (WAB API) + Internal usage (VizieR adapted) using a VI interface (A.Vanhulle)

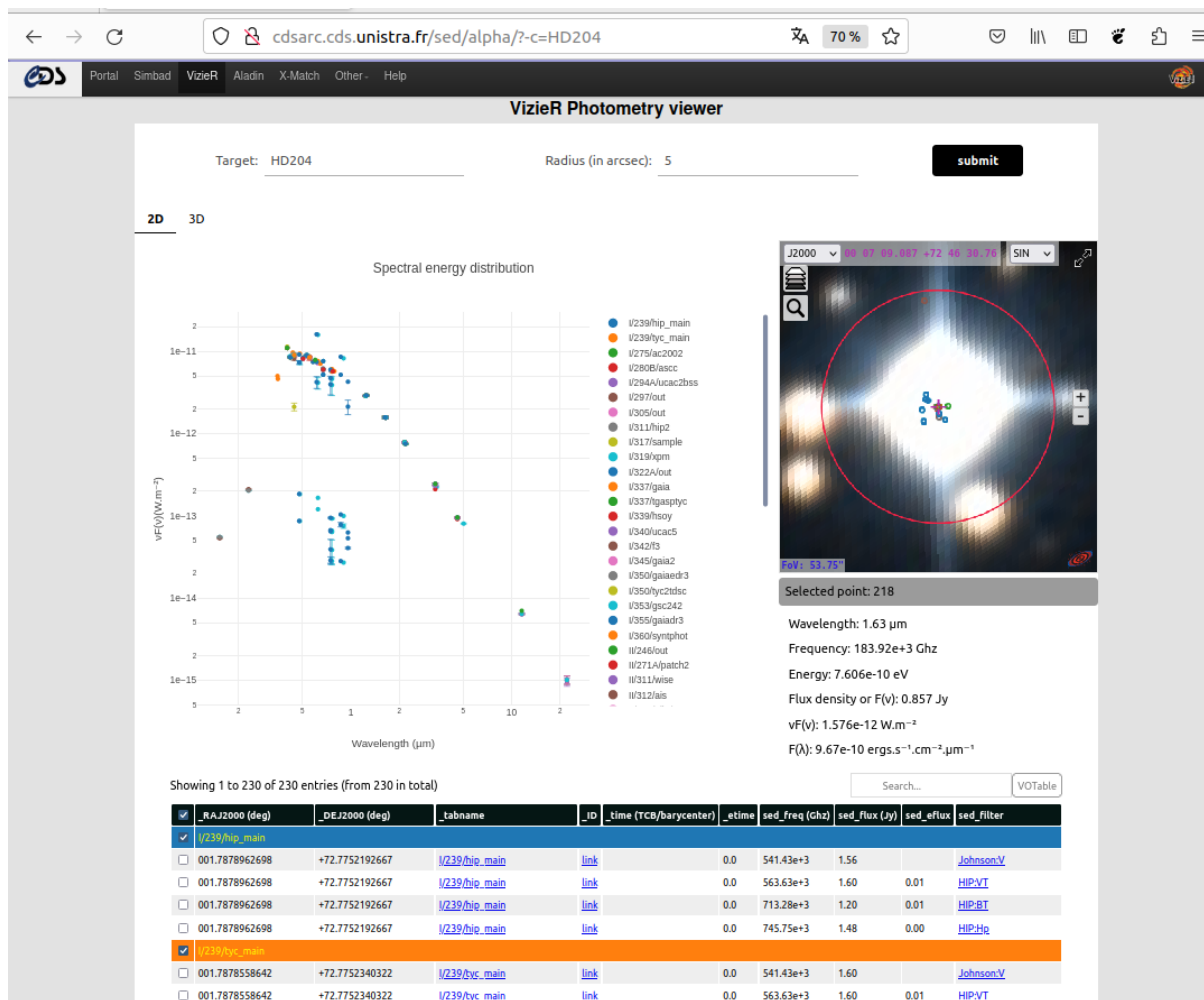
Staaf:
A.Vanhulle, G.Landais, L. Demange,
S.Derriere

```
DEdeg      deg      Declination in decimal degrees (J2000)
>>>>      pos.eq.dec      automatic rule

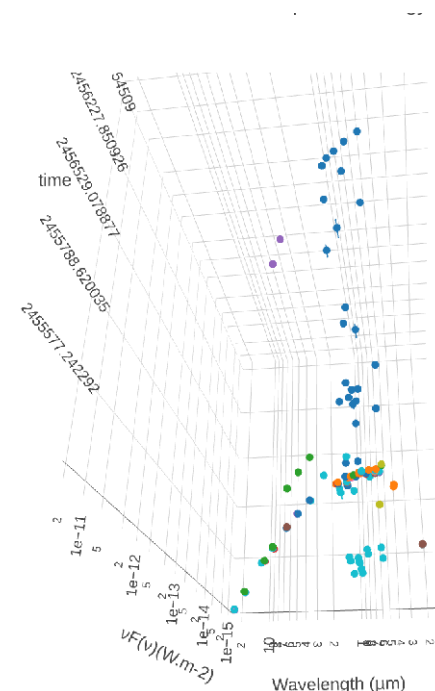
BpRp      mag      Gaia DR3 (Bp-Rp) color
-----   phot.color;em.opt.B;em.opt.R      UCD1 translator (PHOT_CI_B-R)
>>>>>    phot.color      UCD-finder (count in VizieR: 3763)
          src.class.color      UCD-finder (count in VizieR: 51)

Voff      km/s      Relative velocity offset
          phys.veloc;arith.ratio;instr.offset      UCD-finder
          pos.distance;instr.offset;arith.diff      UCD-finder (count in VizieR: 1)
>>>>>    phys.veloc;arith.ratio      UCD-finder (count in VizieR: 9)
          phys.veloc;arith.diff      UCD-finder (count in VizieR: 302)
          pos.distance;arith.ratio      UCD-finder (count in VizieR: 21)
          pos.distance;arith.diff      UCD-finder (count in VizieR: 16)
          phys.veloc      UCD-finder (count in VizieR: 3724)
          instr.offset      UCD-finder (count in VizieR: 225)
          pos.distance      UCD-finder (count in VizieR: 3034)
```

VizieR renewal Technology migration (SED



Staat: Evguenia Sobine,
G.Landais, T.Boch,
A.Vanhulle



Other activities



- Waiting for the **certification** renewal (Core Trust Seal)
(G.Landais + P.Fernique + F.Genova + M.Allen)
- Improve Data dissemination through the **Virtual Observatory**
Data Origin & rich metadata (MIVOT) in in VO output
- Authors **documentation**: tutorial (A.Gonneau, E.Perret, M.Marchand)
*"The journey of your data through the Virtual Observatory
and the European Open Science Cloud"*
<https://cds-astro.github.io/a-FAIR-journey-for-astronomical-data/>
- Notebooks (M.Marchand, G.Landais)
see S. Derriere, M.Marchand demo
- **Name resolver sesame**
Use the **IMCCE** resolver to resolve solar system objets (J.Berthier, J. Normand)
<https://cds.unistra.fr/cgi-bin/nph-sesame/@I?Saturn>

tested in a beta version of Aladin



Outlook – 2024 and (a bit) beyond



- Top Priority = continued support of FAIR data publication from major journals and data producers through variety of access modes
- => continued support and development of in-house tools to assist catalogue indexation and ingestion
- Med to long term goals:
 - Code renewal (Very high priority, ongoing)
 - Consolidation of VizieR team with new staff
- Associated data workf bw: FNSO CDS-ODAS proposal successful
 - 2x2yr (1 astronomer postdoc + 1 dev/infra engineer)
 - Rebuild/improve workf bw for associated data
 - Investigation alternative technologies (e.g. dataverse?)



**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE**

*Liberté
Égalité
Fraternité*

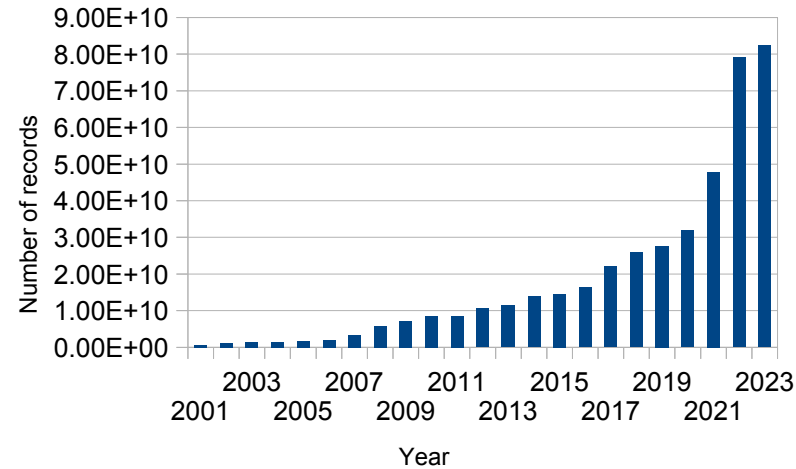
VizieR content - II



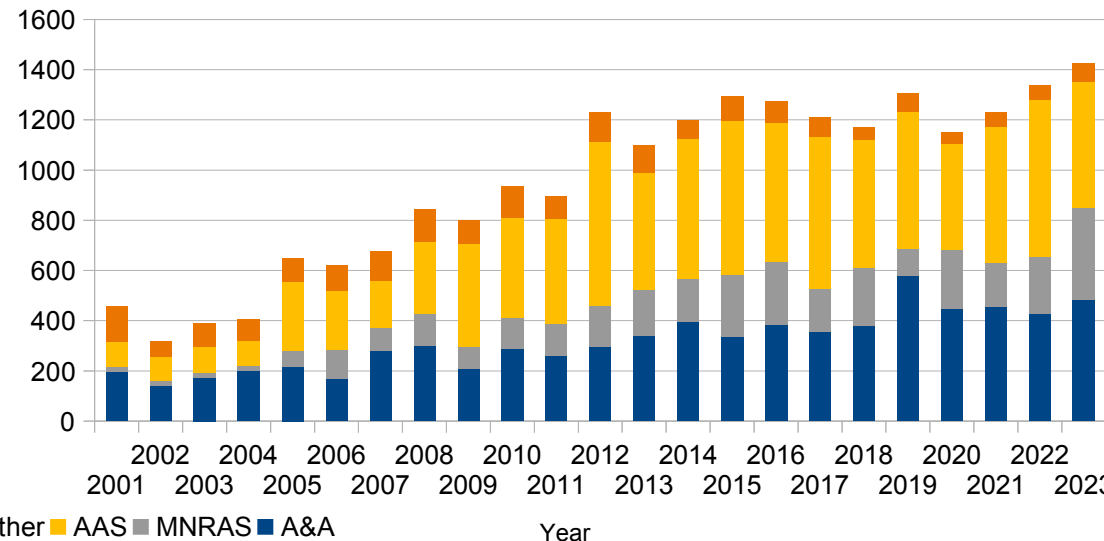
Ingestion statistics :

- A steadily increasing volumetry : +1440 in past year
 - Total (23/10/2023): 24,581 catalogues, >56k tables
- Dramatic increase of records : 80 billions records (+62%), 85.5TB (+60%)

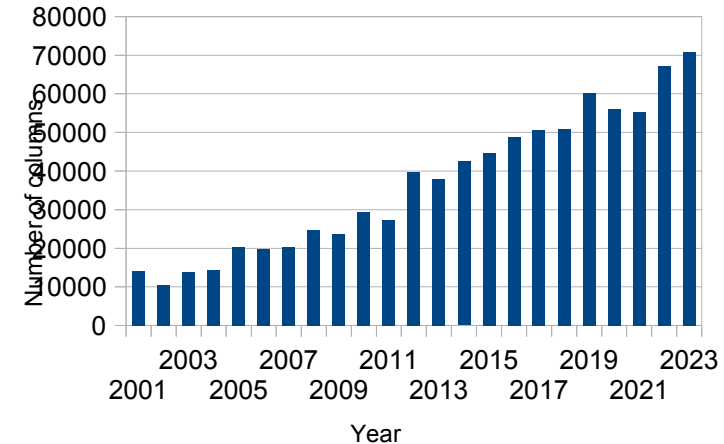
Number of records evolution



Catalogues evolution



Number of columns evolution



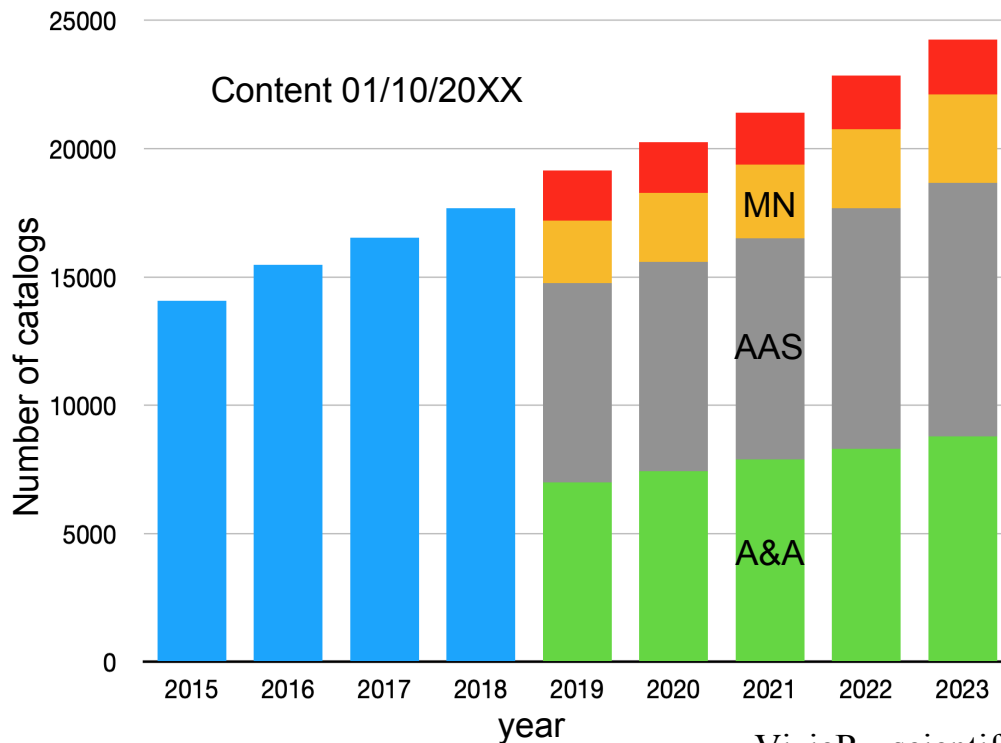
il 2023

VizieR content - II



Ingestion statistics :

- A steadily increasing volumetry : +1419 in past year
 - Total (23/11/2023): 24 483 catalogues, >56k tables
- Slower increase of records (post-Gaia DR3) : ~82 billions records



VizieR - scientific council 2023

