



---

## Scientific council 2017

VizieR Staff and contributors:

Astronomers: P.Ocvirk, C. Bot

Engineers: G.Landais, T.Boch, F.X.Pineau,

Documentalists: P.Vannier, E.Perret,

S.Guéhenneux, M.Brouty, T. Pouvreau.

Postdocs: H. Arab

Non-CDS: L.Michel, J.Y.Hangouet, T.Keller  
(Strasbourg Observatory)



# What is VizieR ? - I



The VizieR **service** provides unified access to a very large collection of astronomical catalogues & associated data:

- Reference catalogues & surveys of astronomical sources at all wavelengths (e.g. GAIA, 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, ...

# What is VizieR ? - II

VizieR can be queried:

- By catalog name
- By wavelength, Mission name, object type / process, ...
- By keyword (galaxies, quasars, ISM, ...)
- By column description
- By position, through ALL catalogs

## Find catalogs among 16056 available

Expand search

**?** *Catalog, author's name, word(s) from title, description, etc. e.g.: AGN, Veron, I1239, or bibcodes...*

- ▶ **Search for catalogs by column descriptions (UCD)** **?**
- ▶ **Search for catalogs containing additional data**

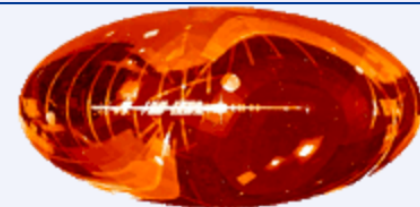
Wavelength	Mission	Astronomy
Radio	AKARI	Abundances
IR	ANS	Ages
optical	ASCA	AGN
UV	BeppoSAX	Associations
EUV	CGRO	Atomic_Data
X-ray	Chandra	Binaries:cataclysmic
Gamma-ray	COBE	Binaries:eclipsing

## Search by Position across 17123 tables


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

Target dimension:

Radius  Box size



**i** [More about VizieR](#)



# Metadata



The metadata describes the data:

- Type of data (catalog, spectra, light curves etc...)
- Units
- Wavelengths / filters / instrument
- Magnitude system, Coordinates system

... with homogenized descriptions (metadata)

Goal: **Discoverability, reusability**, platform for higher level services (phot. Viewer, associated data, ...)

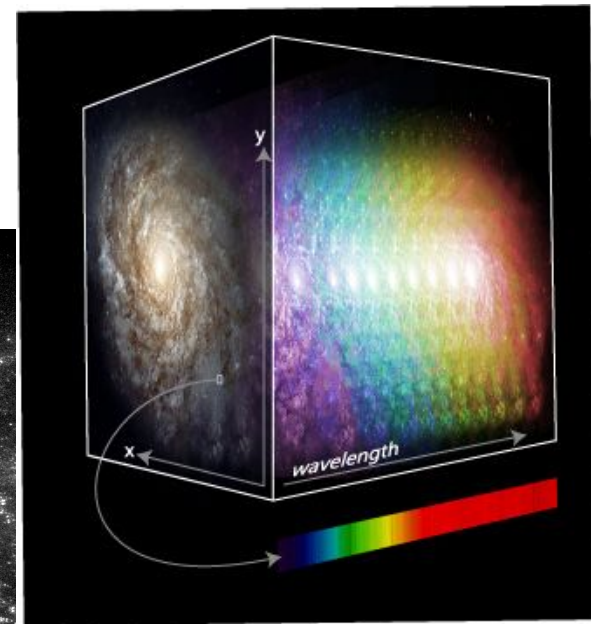
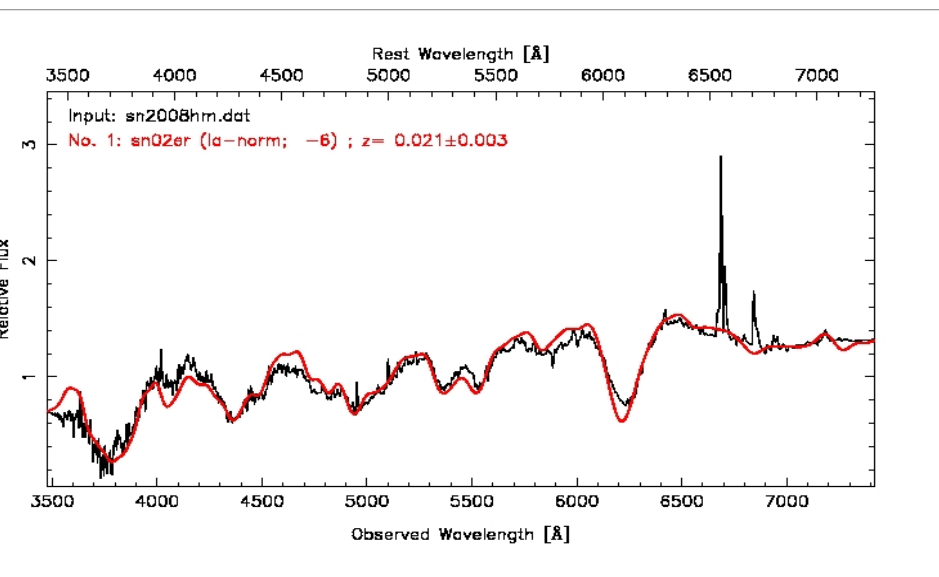
# Associated data

Vizier's content is mostly tabular data (catalogs, line lists...)

More and more, astronomers are required to publish the data that was used to produce the published tables: images, spectra, data cubes, time series...

Same requirements for **discoverability**, **reusability** as tabular data.

Preferably in reduced form rather than raw (available at telescope archive)



# Workflow evolution



- Ingestion of 1 catalogue requires exchanges between up to 3 experts: 1 documentalist, 1 astronomer, 1 software engineer (large cats)
- => Use of project management tools for Ingestion follow-up:
  - Large cats: redmine
  - Small cats: follow-up interface developed in-house (G. Landais)
- Improves traceability and visibility of history of interactions, progress and decisions taken, quality control.

**Ingestion grands catalogue #2846** Mettre à jour Saisir temps Surveiller Copier Supprimer

**PanSTARRS DR1** « Précédent | 1 sur 12 | Suivant »

Ajouté par Thomas Boch il y a 3 mois. Mis à jour il y a 3 mois.

<b>Statut:</b>	Nouveau	<b>Début:</b>	01/07/2017
<b>Priorité:</b>	Normal	<b>Echéance:</b>	
<b>Assigné à:</b>	-	<b>% réalisé:</b>	<div style="width: 8%;"><div style="width: 8%;"></div></div> 8%
<b>Catégorie:</b>	-	<b>Temps passé:</b>	-
<b>Version cible:</b>	-		

[create\\_tables.sql.txt](#) (2,26 ko) Thomas Boch, 07/07/2017 15:35

**Sous-tâches** Ajouter

Ingestion grands catalogue #2847: Téléchargement PanSTARRS DR1	En cours	Thomas Boch	<div style="width: 20%;"><div style="width: 20%;"></div></div>
Ingestion grands catalogue #2848: Conversion en rcf PanSTARRS DR1	Nouveau	Thomas Boch	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2849: Ecriture ReadMe PanSTARRS DR1	Nouveau	Emmanuelle Perret	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2850: Intégration PanSTARRS DR1 dans serveur de test (cdsarc/lo...	Nouveau	Emmanuelle Perret	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2864: Validation scientifique PanSTARRS DR1	Nouveau	Arnaud Siebert	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2865: Publication dans VizieR PanSTARRS DR1	Nouveau	Emmanuelle Perret	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2866: MOC et cartes de densité PanSTARRS DR1	Nouveau	Thomas Boch	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2867: Intégration dans TAPVizieR PanSTARRS DR1	Nouveau	Gilles Landais	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2868: Intégration dans xmatch PanSTARRS DR1	Nouveau	François-Xavier Pineau	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2869: HIPS catalogue PanSTARRS DR1	Nouveau	François-Xavier Pineau	<div style="width: 0%;"><div style="width: 0%;"></div></div>

**Demandes liées** Ajouter

lié à ApJS - Evolution #3178: 2017ApJS...232....4T	Nouveau	29/09/2017	<div style="width: 0%;"><div style="width: 0%;"></div></div>
--	---------	------------	--

# Workflow evolution



- Ingestion of 1 catalogue requires exchanges between up to 3 experts: 1 documentalist, 1 astronomer, 1 software engineer (large cats)
- => Use of project management tools for Ingestion follow-up:
  - Large cats: redmine
  - Small cats: follow-up interface developed in-house (G. Landais)
- Improves traceability and visibility of history of interactions, progress and decisions taken, quality control.

**Ingestion grands catalogue #2846** Mettre à jour Saisir temps Surveiller Copier Supprimer

**PanSTARRS DR1** « Précédent | 1 sur 12 | Suivant »

Ajouté par Thomas Boch il y a 3 mois. Mis à jour il y a 3 mois.

<b>Statut:</b>	Nouveau	<b>Début:</b>	01/07/2017
<b>Priorité:</b>	Normal	<b>Echéance:</b>	
<b>Assigné à:</b>	-	<b>% réalisé:</b>	<div style="width: 8%;"><div style="width: 8%;"></div></div> 8%
<b>Catégorie:</b>	-	<b>Temps passé:</b>	-
<b>Version cible:</b>	-		

[create\\_tables.sql.txt](#) (2,26 ko) Thomas Boch, 07/07/2017 15:35

**Sous-tâches** Ajouter

Ingestion grands catalogue #2847: Téléchargement PanSTARRS DR1	En cours	Thomas Boch	<div style="width: 10%;"><div style="width: 10%;"></div></div>
Ingestion grands catalogue #2848: Conversion en rcf PanSTARRS DR1	Nouveau	Thomas Boch	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2849: Ecriture ReadMe PanSTARRS DR1	Nouveau	Emmanuelle Perret	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2850: Intégration PanSTARRS DR1 dans serveur de test (cdsarc/lo...	Nouveau	Emmanuelle Perret	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2864: Validation scientifique PanSTARRS DR1	Nouveau	Arnaud Siebert	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2865: Publication dans VizieR PanSTARRS DR1	Nouveau	Emmanuelle Perret	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2866: MOC et cartes de densité PanSTARRS DR1	Nouveau	Thomas Boch	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2867: Intégration dans TAPVizieR PanSTARRS DR1	Nouveau	Gilles Landais	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2868: Intégration dans xmatch PanSTARRS DR1	Nouveau	François-Xavier Pineau	<div style="width: 0%;"><div style="width: 0%;"></div></div>
Ingestion grands catalogue #2869: HIPS catalogue PanSTARRS DR1	Nouveau	François-Xavier Pineau	<div style="width: 0%;"><div style="width: 0%;"></div></div>

**Demandes liées** Ajouter

lié à ApJS - Evolution #3178: 2017ApJS...232....4T	Nouveau	29/09/2017	<div style="width: 0%;"><div style="width: 0%;"></div></div>
--	---------	------------	--

# Followup application



The service enables to :

- Choose and notify the astronomer (automatically) in function of configurable criteria
- To validate/refuse the catalogue availability with notification to the documentalist in charge of the catalogue (notification by mail and highlighted in the web application)

VizieR ingestion followup

Only ready cats.  Catalogue

Show 10 entries Search:

status	name	title	doc	validator	release	comment	action
wait	JApJS/230/3 meta	Optical & Spitzer photometry in IC 1805 (Sung+, 2017)		Caroline Bot	2017-11-02 09h59	cf mail Pierre	<input checked="" type="checkbox"/>
test	I/335 meta	GALEX-GR6/7 data release (Bianchi+ 2014)		Pierre Ocirk	2018-05-31 00h00		<input checked="" type="checkbox"/>
test	JApJ/797/116 meta	Chemical properties of M31 star clusters (Colucci+, 2014)	mb	Pierre Ocirk	2017-10-27 13h58		<input checked="" type="checkbox"/>
test	JApJS/229/28 meta	SFINCS: X-ray & IR catalogs & membership (Getman+, 2017)	ep	Caroline Bot	2017-10-19 13h22	-table 1: un probleme de formattage visiblement: -table pbmmbrs: les liens simbad basés sur les identificateurs 2MASS ne marchent pas	<input checked="" type="checkbox"/>
test	JApJ/834/52 meta	The Megamaser Cosmology Project (MCP). IX. (Gao+, 2017)		Pierre Ocirk	2017-09-29 09h16		<input checked="" type="checkbox"/>
test	JApJ/834/122 meta	Massive star formation in the LMC. I. N159 & N160 (Gordon+, 2017)		Caroline Bot	2017-09-29 09h14		<input checked="" type="checkbox"/>
readyprod	J/A+A/606/A21 meta	LMC bar star clusters (Piatti, 2017)		Pierre Ocirk	2017-09-29 08h42		<input checked="" type="checkbox"/>
readyprod	J/A+A/606/A18 meta	WASP-103b light curves (Lendi+, 2017)		Pierre Ocirk	2017-09-28 09h54		<input checked="" type="checkbox"/>
readyprod	J/A+A/606/A13 meta	QSO eHAQ0111+0641 spectra (Fynbo+, 2017)		Pierre Ocirk	2017-09-27 10h08		<input checked="" type="checkbox"/>
readyprod	J/A+A/606/A12 meta	MUSE-Wide survey: 831 emission line galaxies (Herenz+, 2017)		Pierre Ocirk	2017-09-26 11h08		<input checked="" type="checkbox"/>

Showing 1 to 10 of 63 entries

Previous 1 2 3 4 5 6 7 Next

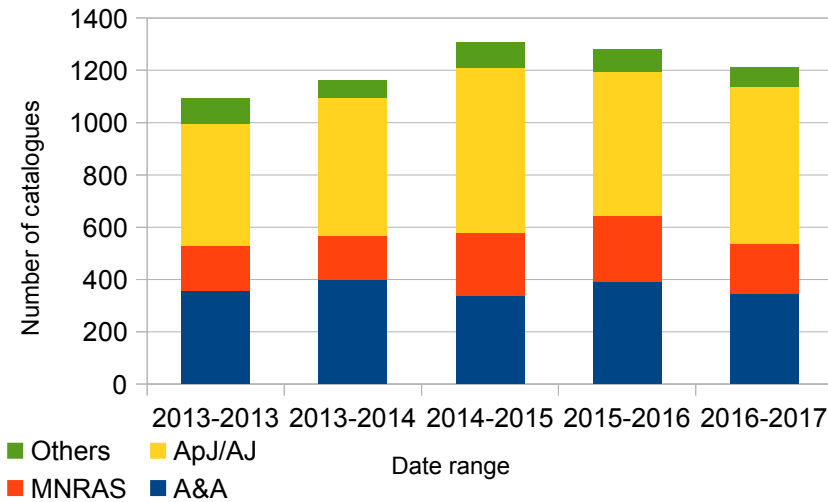




## Statistics in input

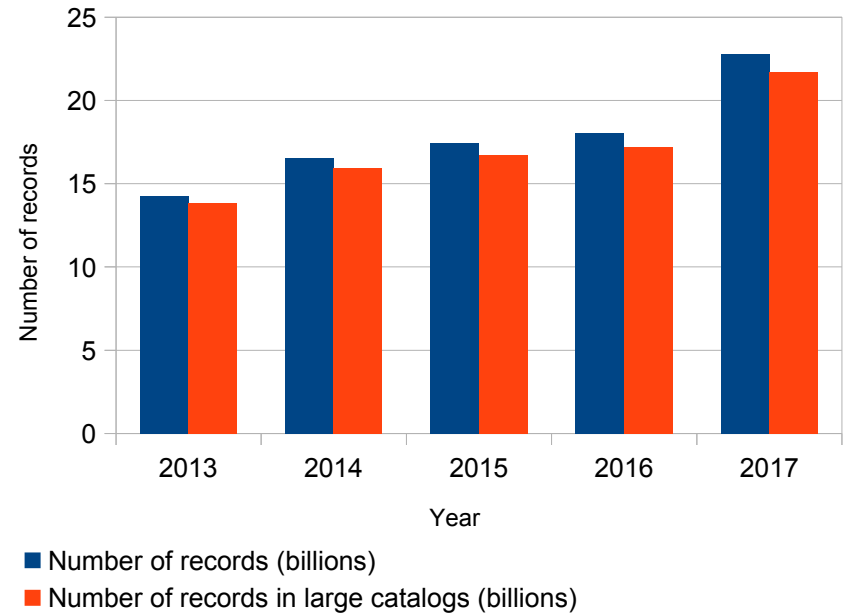
- A steadily increasing volumetry
  - 16,480 catalogues, 35,593 tables, 22,8G records

VizieR catalogue ingestion



Evolution of records available in VizieR

(from september to september)



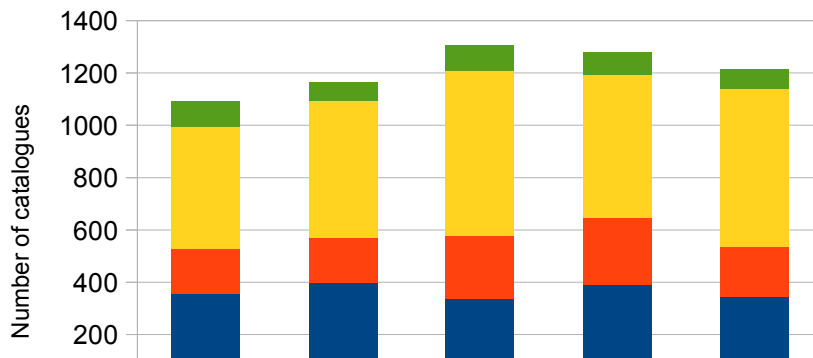
# Large cats



## slides

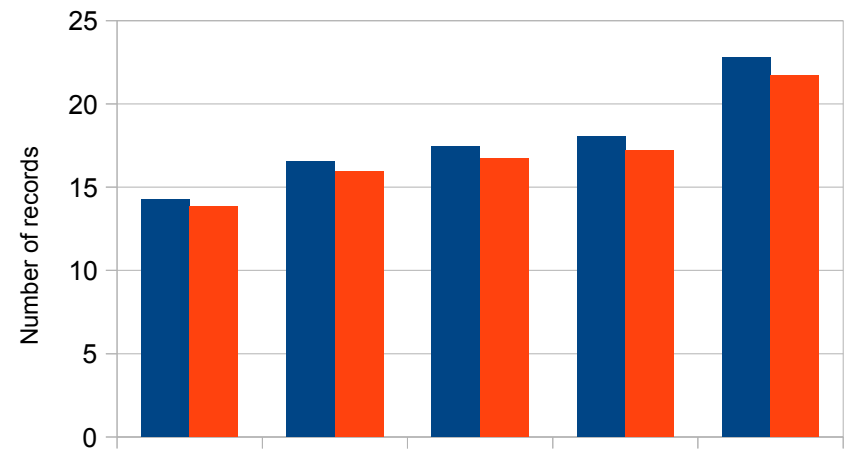
- 1, 3, 4 from Thomas
- 2 from Thomas,
- Slide about followup application
- About stats : 2017 : formation tiphaine, baisse MNRAS, increase in quality ?
  - 16,480 catalogues, 35,593 tables, 22,8G records

VizieR catalogue ingestion



Evolution of records available in VizieR

(from september to september)




# New large catalogues in VizieR



10 large catalogues ingested (oct. 2016 —> sept. 2017)

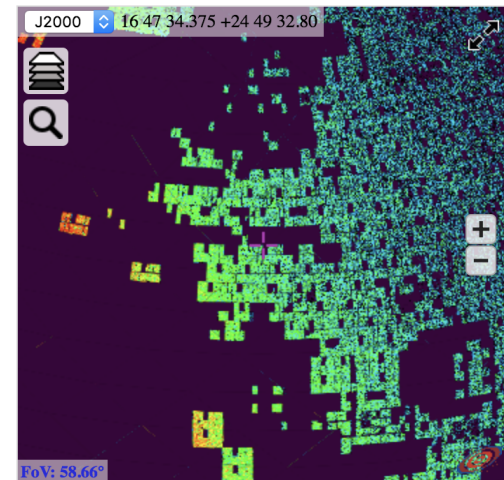
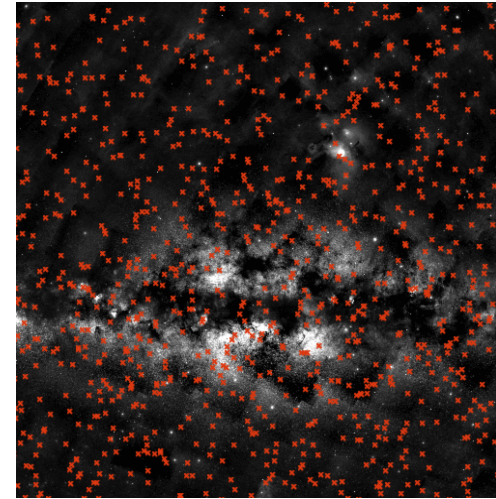
*Total: 4 billion rows, 1.5 million queries*

- SDSS DR12
- Hubble Source Catalog (DR1 and DR2)
- UCAC5
- HSOY
- K2 Epic
- ESO [public surveys](#) catalogues 
  - KiDS DR2 and DR3
  - Viking DR2
  - VVV DR2
- Chandra galactic bulge survey
- Expected in 2018: Pan-STARRS DR1 full, GPS1 (Gaia-PanSTARRS-SDSS proper motions), GAIA DR2, ...

# HiPS catalogues and HiPS density maps



- HiPS catalogues
  - 15 newly created: *Viking DR2*, *UCAC5*, *HSOY*, *SPM4*, *APASS9*, *KIDS DR2*, *Tycho 2*, *2MASS 6X*, *IGSL3*, *GSC 2.2*, *GSC 1.2*, *PPMX*, *SDSS7*, *Attitude*, *URAT1*
  - Mid-term goal: generate the HiPS catalogues for every VizieR table with positions
- HiPS density maps
  - available for all VizieR tables
  - to be integrated in HiPS clients (Aladin Lite, Aladin Desktop)



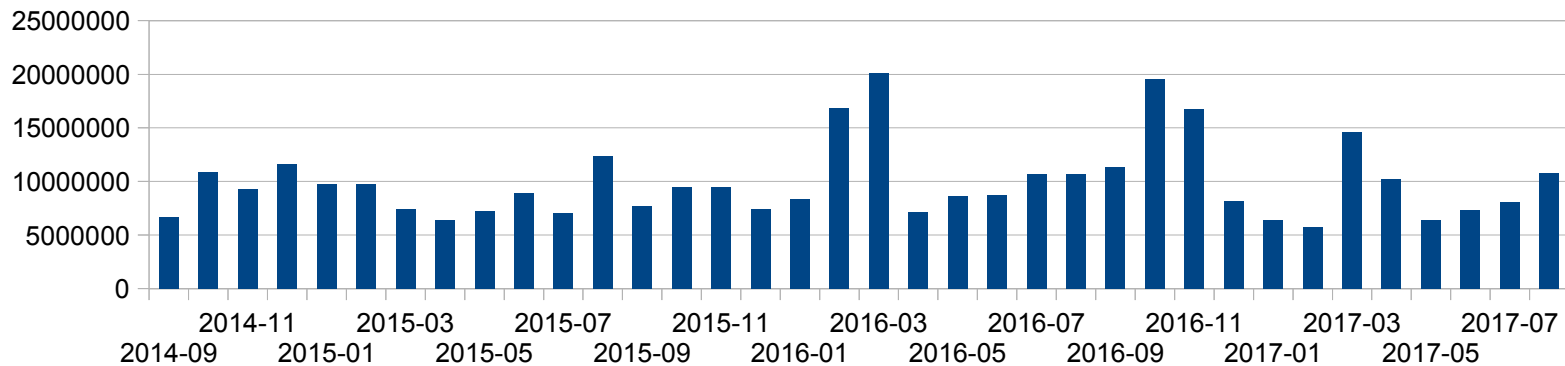
# Output statistics



## A continuous usage of the service

~326,000 queries/day (without bots)

VizieR queries history from sept. 2014

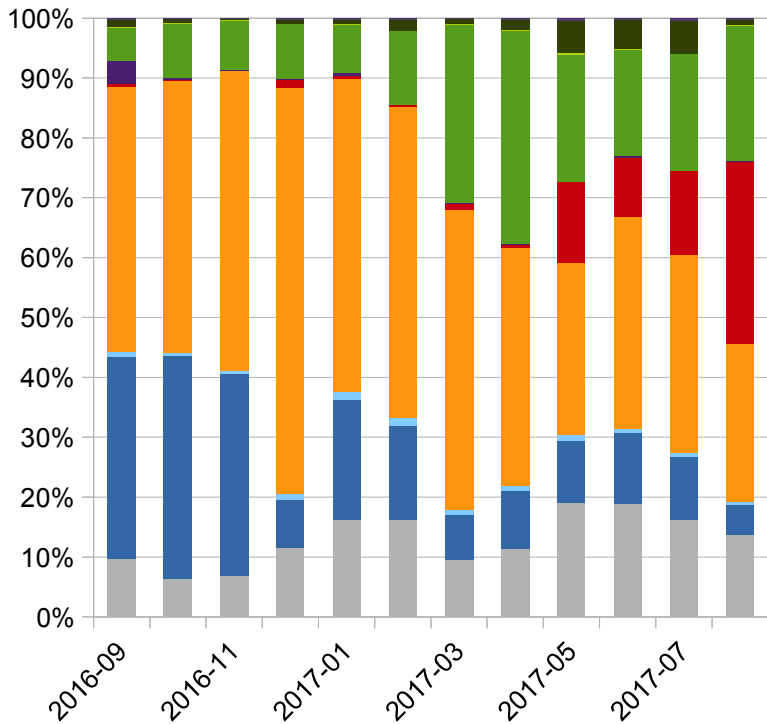


- The **VO - TAPVizieR** access query ing VizieR tables using ADQL/SQL  
~5,200 queries/day (pics in april/may)
- The **new associated data service**  
Search spectra/images in FITS format in VizieR  
~270 queries/day (pics in January/Februray)

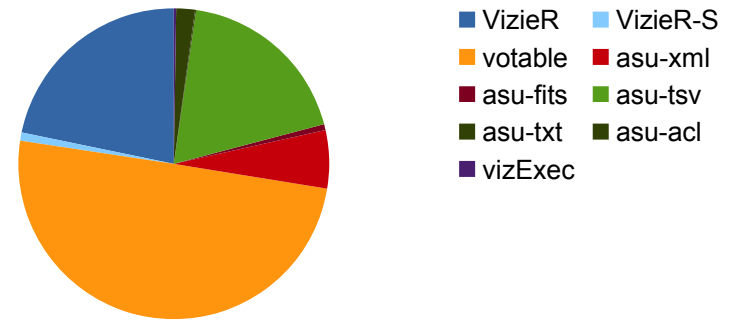
# Output popularity



VizieR queries repartition  
(sept 2016 - sept 2017)



VizieR Queries repartition  
(sept 2016 - sept 2017)



**VOTable** 50.2%

- Python: astroquery
- TOPCAT (JAVA)
- Aladin (JAVA)

**Asu-XML** 6.2%

- Python
- Aladin

**TSV** 18.6%

- Wget, curl



- Faster access for local users
- Data redundancy
- Mirrors are used as backups during CDS downtimes (automatic switch)
- Mirror status
  - Updated on a regular basis
  - Hawaii mirror is OFF
  - UK mirror was OFF during ~ 1 month
  - CfA mirror is available and is waiting for a migration

# The associated data service



## To search spectra/images in VizieR

- The VizieR indexation process to search FITS spectra/images
- A VO-service with meta-data using the ObsCore DataModel of the Virtual Observatory and with access service: SIA/SSA, ObsTAP

Search associated data among the VizieR catalogues

This web page is an access to the VizieR Associated data (images, spectra, timeseries, SED) which comes from publications. This tool is the result of the documentation assigned by the authors of the catalogues and supervised by the CDS documentalists team (see the VizieR ingestion tool).

**VO compatibility**  
The meta-data and the search engine are built according to the VO framework (SIA, SSA, ObsTAP) and can so be queried by VO softwares. The data are gathered with the Saada engines, and the VO data model ObsCore has been chosen for the documentation.

Simple search  ObsTAP Query

Q Search by position : 266.416833 -29.007806 radius 1 deg

Q Search by spectral band : min max  $\mu\text{m}$

Q Search by time data : start stop (MJD)

Q Search by catalog/identifier:

Spectrum  Time series  Image

500 entries max

55 entries

Preview	Target	Data collection	Ra	Dec	Band min (nm)	Band max (nm)	Begin time (MJD)	End time (MJD)	Facility
	SCP05M24	II/243	266.671	-29.018	7,000,000	15,000,000	50,872.691		ISO
		JJA+A/515/A42	266.745	-28.503	3,600,000	4,500,000			Spitzer
		JApJS/175/277	265.828	-28.502	100,000	10,000,000	51,544.500		JCMT

## The meta-data assignment process

The engine uses the Saada database generator to make the mapping with ObsCore DM

- With an automatized mapping using the FITS headers
- Completed with human assignment



# The associated data service



## The Contents

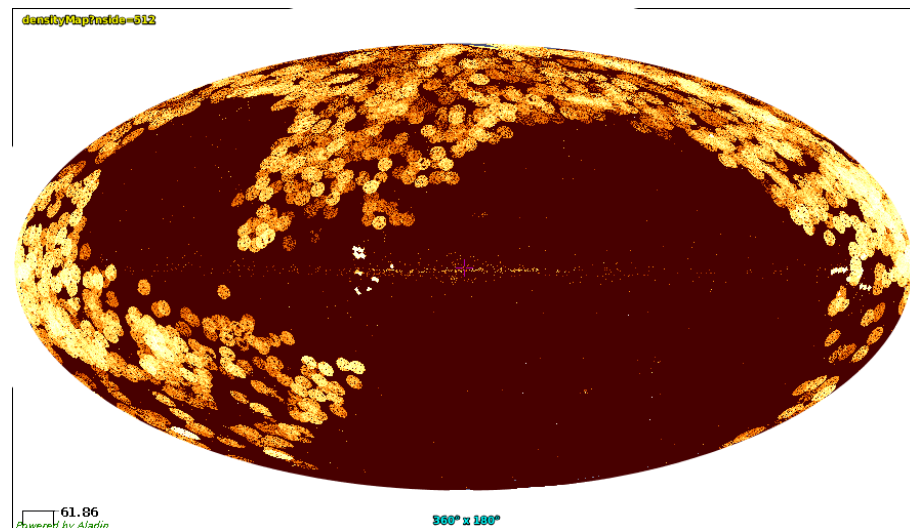
Ingestion of ~140 VizieR catalogues having FITS

### Images content

- 31,237 images from 84 catalogues
- largest collection:
  - ATLASGAL: 9K images (Infra red)
  - S4G: 8.4K (Infra red)

### Spectra/time-series content

- 62 catalogues :
  - ~102,000 spectra coming from various publications
  - + 170,000 time series from CoRoT
  - + 2.2M spectra from LAMOST DRA and 4.1M spectra from LAMOST DR2



# The associated data service



## Ingestion pipeline is available from end to end

- **VizieR associated data access (VO access + dedicated web page) is available since 2016**
- The new web **VizieR submit service** is available for authors in beta version since january 2017
- The **xfits** application used in the CDS for documentalists which enable to map FITS header with the ObsCore DM had been updated wit preview and ingestion job submission

You are logged as testgilles ✕ [report](#)

### VizieR catalogue upload

Upload tabular data    Fill the README    Upload Spectra/TimeSeries    Upload Images    Terminate

**i** Upload spectra/time-series in VizieR and provide them through a dedicated database. Providing these documents need description for indexation. Currently, the indexation is available only for FITS document The VizieR engine will first extract the metadata from the documents uploaded in a mapping that you can update or change. [Next step](#)

#### Upload your spectra/ time series

You have some spectra/ time series **1** Only FITS format is accepted! Please, upload documents in other format later.

You can upload your documents one by one by describing them independently  
OR if you have documents with similar header **2** you can upload a collection (an archive in tar, zip format) and put a common description.

Add new document(s) [Browse...](#) No file selected. [Upload](#)

✓ [File spectrum/J\\_AA\\_414\\_699\\_hd20766.fits](#) (10 item(s) filled)

Force my mapping

Target name	<input type="text" value="OBJECT"/>	Assigned by Keyword <b>1</b> Value: hd 20766
Right ascension	<input type="text" value="RA"/> <input type="text" value="ICRS"/>	Assigned by Keyword <b>1</b> Value: 49.47916666666664
Declination	<input type="text" value="DEC"/>	Assigned by Keyword <b>1</b>

# The associated data service



## The human meta data assignment effort

- Proper indexation is the responsibility of authors and documentalists
  - The new submit web interface (beta version) had been chosen by ~40 authors, 11 having FITS
  - authors need to be educated (communication effort planned)
- Data complexity + increased indexation workload still in learning period  
=> currently only a fraction of the incoming data is processed.
- => Existing catalogs ingestion executed by H.Arab (postdoc)  
+ G.Landais (software engineer)

# Other tasks accomplished



## Ingestion tools

**The XML extractor tools** « getapj » (T.Delacour) dedicated for AAS journals (AJ, ApJ, ApJS) is used by documentalists.

The software facilitates the documentation work :

- downloads tables
- extracts keywords, abstract, reference
- prepare a ReadMe skeleton

## Maintenance

- TAPVizieR migration on a duplicated architecture
  - PostgreSQL version migration + Data migration (10.1Tb)
  - Data and software duplication with front end balancing capability
- VizieR in Docker used locally for TAPVizieR and scheduled for the new CfA mirror

# Tasks Work accomplished



## Organisation

- VizieR certified by the Data Seal of Approval :
  - A certification to renew every 3years
  - New criteria with more importance of data reusability
  - The documentation is ready to submit to the DSA evaluation
- Workflow to notify the catalogues status ingestion had been review
  - A new “Followup” application had been developped this year
  - Completes the “redmine” software used by documentalists
  - A web application with dedicated interfaces for astronomers and documentalists
  - Includes mail notifications and is included in the VizieR ingestion tools





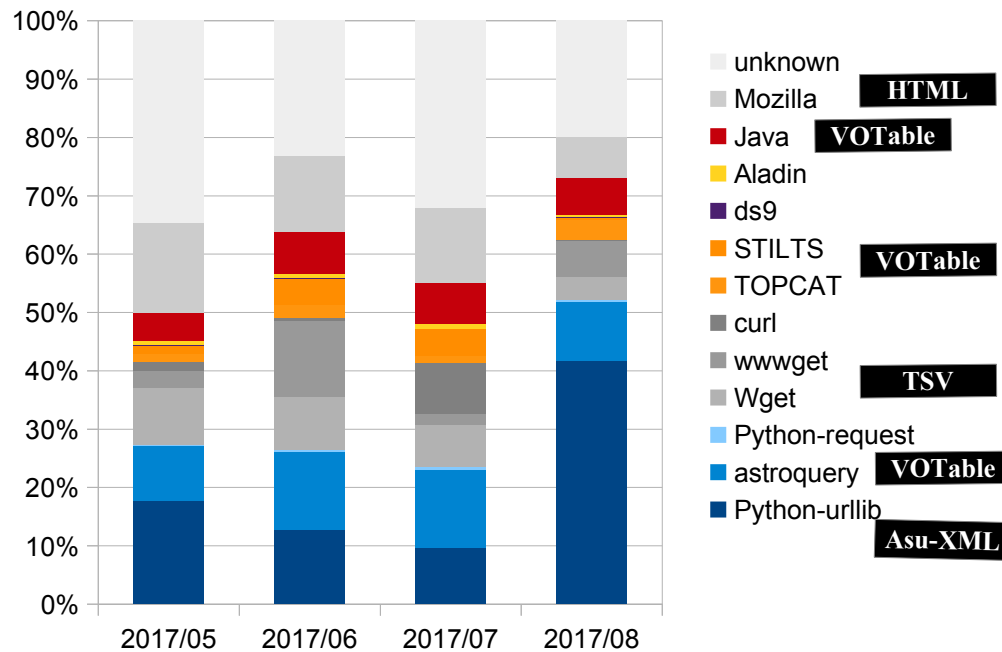
- Put the new beta vizier submission interface to production
- To finalize the Data Seal of Approval certification
- DOI attribution for the VizieR catalogues (discussion during the LISA conference)
- Investigation of meta-data for time domain in VizieR and in the Virtual Observatory

with :

- T.Boch (engineer)
- S.Derriere (astronomer)
- A.Nebot (astronomer) : chair of the IVOA time interest group

## Most represented software requesting VizieR:

VizieR queries per agent



**Astroquery : 366 IP**  
**Python urllib: 77 IP**

### August 2017

**Astroquery** : 366 IP  
 ~1.1M queries from 366 IP in VOTable

**Python urllib**: 77 IP  
 3.8M queries in asu-XML + 700K in TSV

**Wget/wwwget**: 232IP,  
 1.1M in TSV

**JAVA** :404 IP,  
 666K in VOTable

**TOPCAT**: 352 IP,  
 370K in Votable

**Aladin**:4031 IP  
 27K in VOTable+10K asu-XML

**Mozilla**:16K IP,  
 470K in HTML+10K in VOTable