CDS Scientific Council Meeting: Nov 29-30, 2021
Hybrid - *in-person + virtual meeting*

Observatoire astronomique de Strasbourg - Amphitheatre
11 rue de l'Université, Strasbourg

14h Strasbourg, 13h London, 8h Wash. DC

Monday 29 November

14h - Welcome and introductions.
14h10 - CDS Activities 2020-2021. *(M. Allen)*
14h30 - CDS Information System. (P. Fernique)
14h45 - SIMBAD and Biblio. (C. Loup, A. Oberto, S. Lesteven)
15h00 - VizieR. (P. Ocvirk, G. Landais)
15h15 - Aladin. (C. Bot, T. Boch)
15h30 - *Coffee Break*
16h00 - X-Match. (F-X. Pineau)
16h15 - R&D. (A. Schaaff)
16h30 - CDS Science Team Work. (B. Vollmer)
16h50 - Demonstration. (S. Derriere)
17h - Discussion. *(All)*
17h30 - Close.

Tuesday 30 November

10h - 12h : Visits / demos for in-person participants
10h - Demonstration of the Coronelli Globe with data projection (S. Derriere)
- other informal demonstrations

12h - Lunch with all the CDS (University Restaurant 'Le 32')

14h - CDS strategy, plans and challenges. *(M. Allen)*
14h30 - **Closed session**
16h - Close

In-person + Remote Attendees:

(All CDS staff in the open sessions)

Stephen Serjeant [Chair] (Open University, UK) - *in-person*
Eric Peng (Peking University) - *videocon*
Guido De Marchi (ESA) - *in-person*
Michael Sterzik (ESO). - *vidoecon*
Valerie Connaughton (NASA) - *videocon*
Denis Veynante (CNRS, HPC and data) - *apologies*
Philippe Laudet (CNES) - *videocon*
Thierry Forveille (IPAG) - *in-person*
Franck Le Petit (Obervatoire de Paris) - *in-person*
Chiara Ferrari (OCA) - *videocon*
Marian Douspis (IAS) - *videocon*

Bruno Bezard (INSU representative delegated by G. Perrin) - *in-person*
Pierre-Alain Duc (Dir. Obs. Strasbourg) - *in-person*
Representative of Université de Strasbourg - Julie Thompson

CDS Activity Report 2020-2021

November 29, 2021

Mark Allen - Director CDS

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

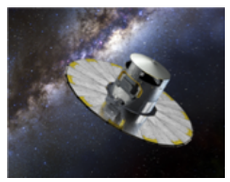
CDS mission

- **Collect useful data on objects in electronic form**
- **Improve them by critical evaluation and combination**
- **Distribute the results to the international community**
- **Conduct research using the data**

Science Driven:

- *Necessary evolutions to meet the scientific reference service needs of the astronomy community*
- *Innovations to meet challenges and ensure sustainability*
  - *Science is changing, technology is changing*

Contributing to the global astronomy data infrastructure



Ground and Space Observatories, Instruments and missions

ESO

esa

cnes

NASA

CDS

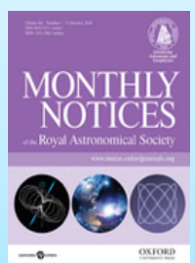
cnrs INSU Observer & comprendre

Université de Strasbourg

Journals

Astronomy & Astrophysics

AAS



+ …

Astronomy Data Centres

astrophysics data system

NED

CADC, MAST, HEASARC, IPAC, + …

Virtual Observatory

EURO VO

France

IVOA

ASOV

H2020 projects:

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

eurPLANET 2024 Research Infrastructure

EOSC

Data e-Infrastructures

RDA
RESEARCH DATA ALLIANCE
+ RDA Europe

CORE TRUST SEAL

WORLD DATA SYSTEM

2020-2021

- **A challenging year that required adaptations to the way of working** — *Re-grouping since September 2021*

- **Core work of building CDS content, and operating the services, was maintained at a very high level.**

- **Improvements to the Information System at various levels, and planning for the future.**

- **Developments to support internal CDS processes, & developments in response to new scientific needs - e.g. time domain!**

2020-2021

Special things in this period:

- Working from home: 30 Oct. 2020 - 02 Sept. 2021.

  (50% in-person in Summer 2021)

- Emphasis on operational stability and continuity.
- Continually adapting procedures to deal with the uncertainty.
- Communications - videocons, rocket-chat.
- Virtual meetings / conferences / training.

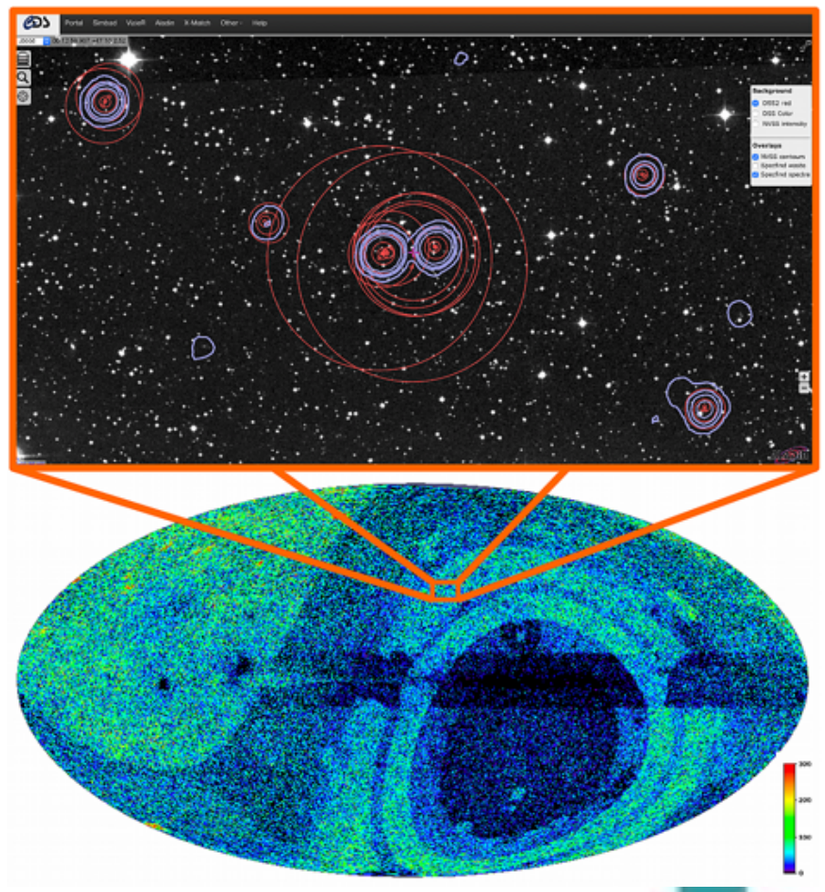
2020-2021

- **CDS services continue to be heavily used**
  - ***1.9 million queries/day*** *- see next presentation*
- Contributions: IVOA, EOSC, RDA, IAU
- Projects — **ESCAPE, EuroPlanet, EOSC-Future, XMM2ATHENA**
- Many community interactions - **AAS, ADASS, EAS**, **LISA**, UNISTRA OAW, French, Chinese, German communities, Time Domain Community, Radio Ast. Community +…
- Engagements with large projects in various ways…
- CDS Science team progress *- see presentation*
- Involvement in networks at national level

Highlights

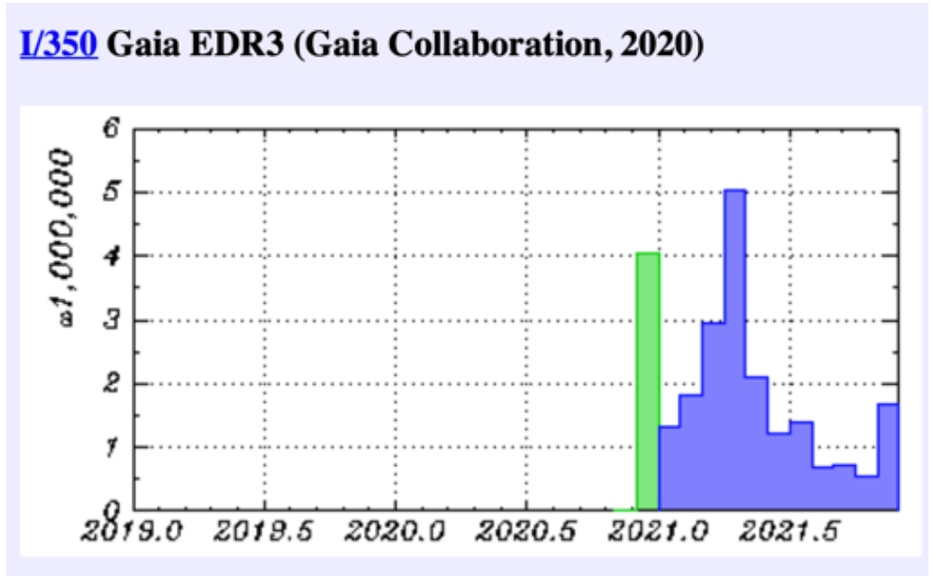
- **New CDS ‘publishing registry’ -** makes CDS services accessible in many tools / services. *(June 2021)*
  - Complies with most recent IVOA registry recommendations
  - Vizier mirrors, MOC footprints, spectral bands, catalogue DOIs, keyword mapping to Unified Astronomy Thesaurus

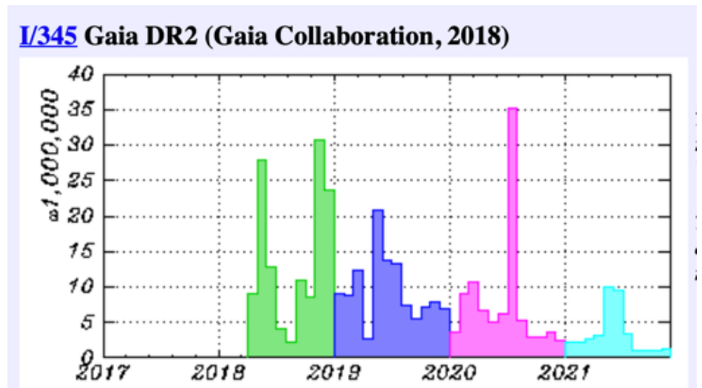
- **SPECFIND V3.0 released** *(October 2021)*
  - Radio continuum source catalogue
  - 204 input tables from Vizier
  - Cross-identified 1.6 M radio sources
  - 340 000 radio spectra created
  - Stein et al. 2021 (A&A 655, A17)
  - ‘Emblematic CDS paper’ :
    - Astronomers, Documentalists, Engineers



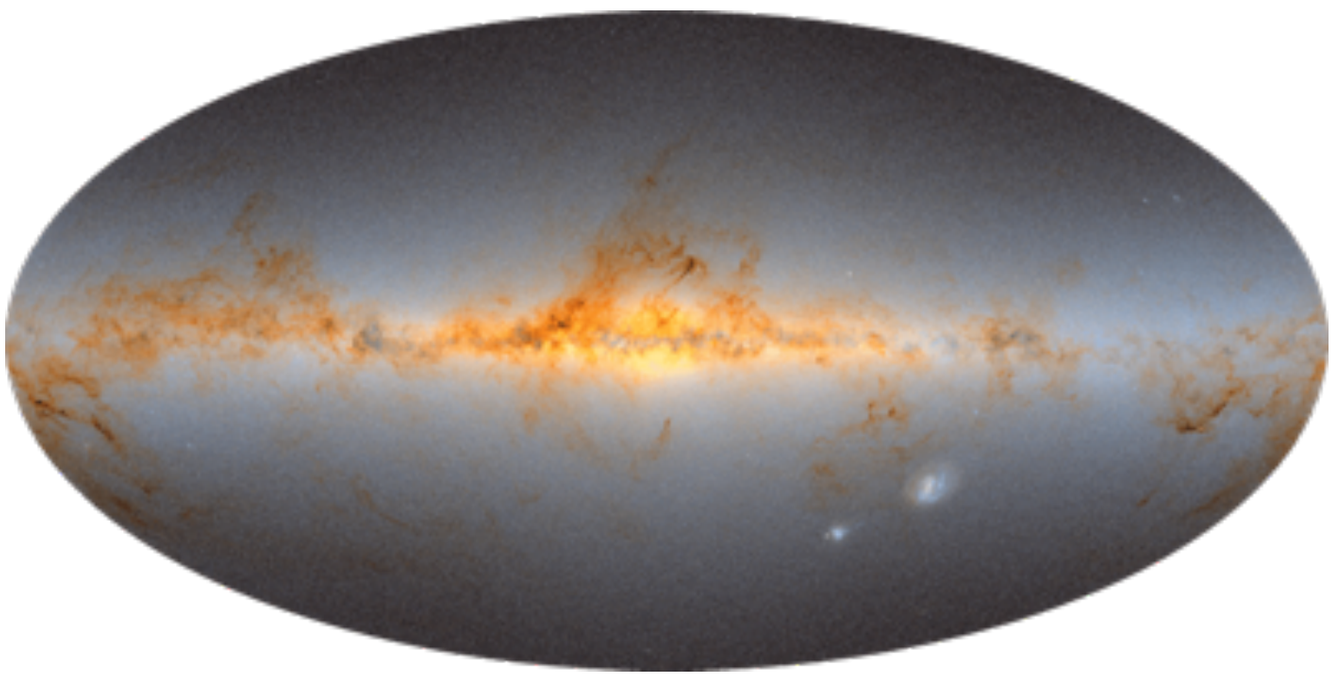
Highlights

- **Gaia Early Data Release 3**
- **3 December 2020**
- 1.8 billion records
- 1.5 billion stars:
  - parallaxes, proper motions, colours
- VizieR, X-Match, Aladin/Aladin Lite, +VO
- CDS generated all-sky flux map, + more
- 4 million queries in 1st month





DR2 also continues to be heavily used



partnerships: ESA/DPAC, CNES

CDS in the community

Astronomical Data Analysis Software and Systems (ADASS) Conference, Virtual (Hosted by IAA-CSIC, Granada), 8-11 November 2020
Invited talk: *Data interoperability – The CDS experience. (Pierre Fernique and the CDS team)*
(see the CDS-ADASS document for all CDS contributions).

New Year Lectures 2021 organized by the NAOC (China National Astronomical Data Center), **18 January 2021**. *SIMBAD, VizieR and Aladin: The CDS Astronomical Suite - P. Fernique*

LINEA workshop On the Future of Data Centers and eScience Institutes, 13-15 April 2021
Invited talk: *CDS services for reference astronomy data - supporting Open Science,* M. Allen.

Library and Information Services in Astronomy (LISA IX) Conference, June 14-18, 2021
Contributed talk: *Impact of Astronomy Evolution on the Documentalists' Activities at CDS*
E. Son, S. Lesteven, F. Marquis, F. Genova, and M. Allen (video link), + posters

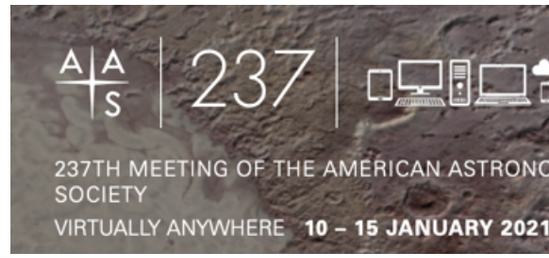
American Astronomical Society (AAS) summer meeting, Virtual, 7-9 June 2021
Special session — Astronomical Data Visualization in the Age of Science Platform
Invited talk: *Aladin Lite, ipyaladin and the HiPS ecosystem,* T. Boch and the Aladin team

European Astronomical Society annual meeting. 28 June 28 - 02 July 2, 2021
Virtual Exhibit - S. Derriere, M. Allen, S. Amodeo, M. Buga, E. Collas, H. Heinl
Special Session - Legacy and stewardship of astronomical archives towards multi-instrument, multi-wavelength and multi-messenger science.
Invited talk — *Spatio-temporal exploitation of astronomical archives,* S. Derriere.

CDS in the community

- **CDS virtual booth** at AAS 237
- ~142 booth visits
- Presentations in zoom
- Community feedback and questions
- Lessons learned
  - Register webinars for better visibility in the program

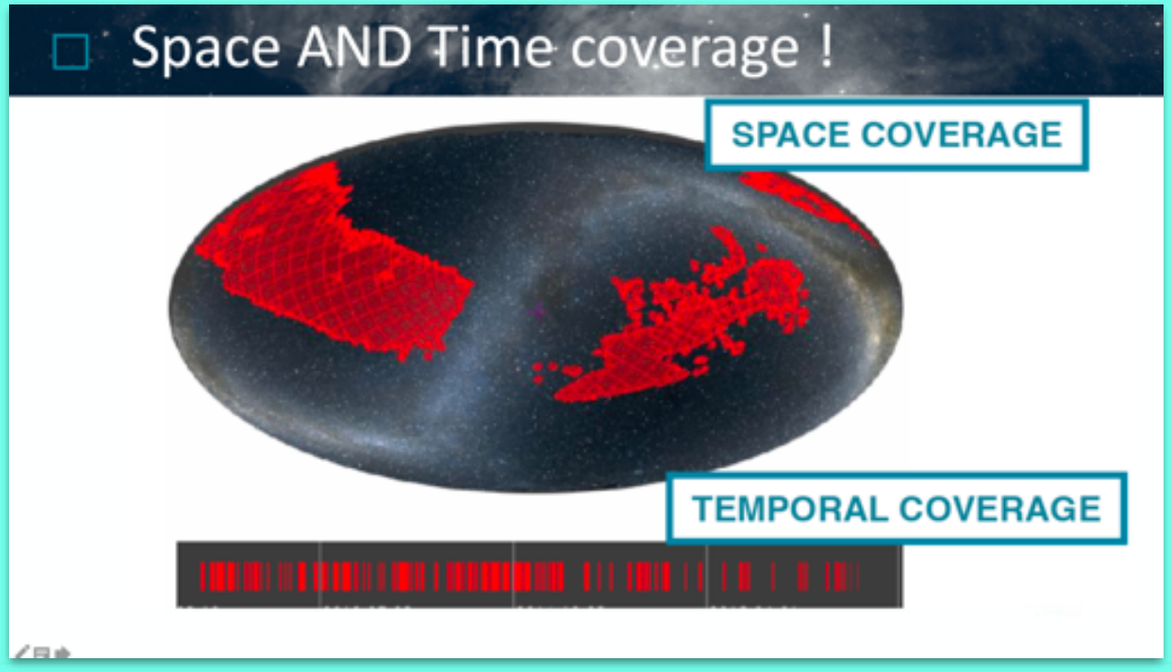
CDS @

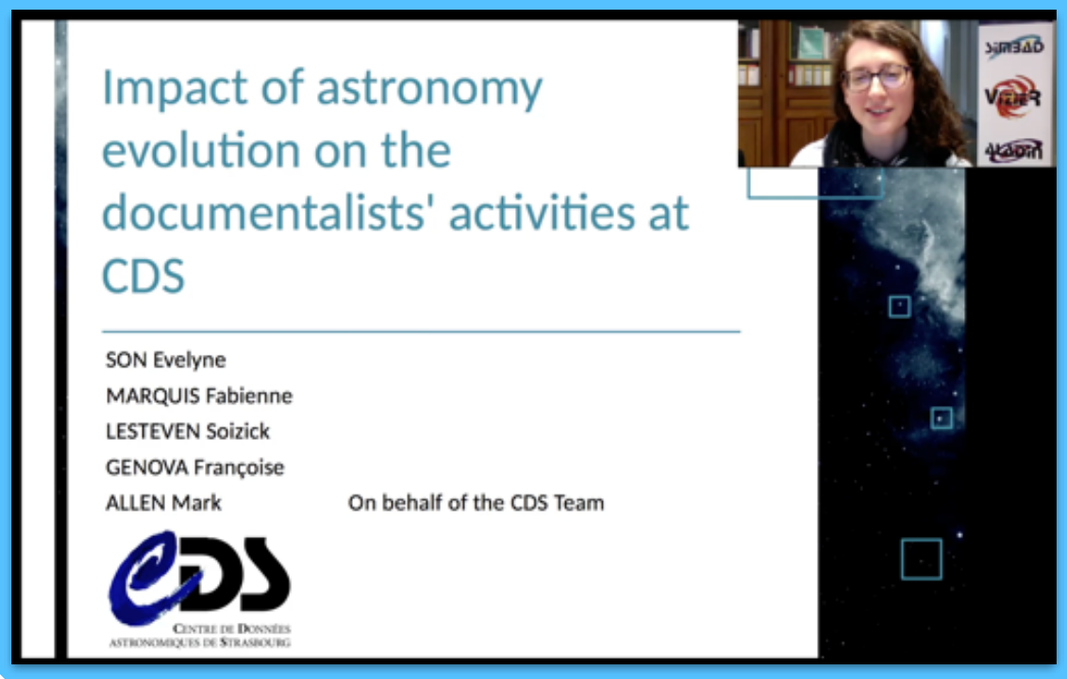


Communication themes:

Library and Information Services in Astronomy (LISA IX) Conference, June 14-18, 2021

Contributed talk: ***Impact of Astronomy Evolution on the Documentalists' Activities at CDS*** - E. Son, S. Lesteven, F. Marquis, F. Genova, and M. Allen





European Astronomical Society annual meeting. 28 June 28 - 02 July 2, 2021

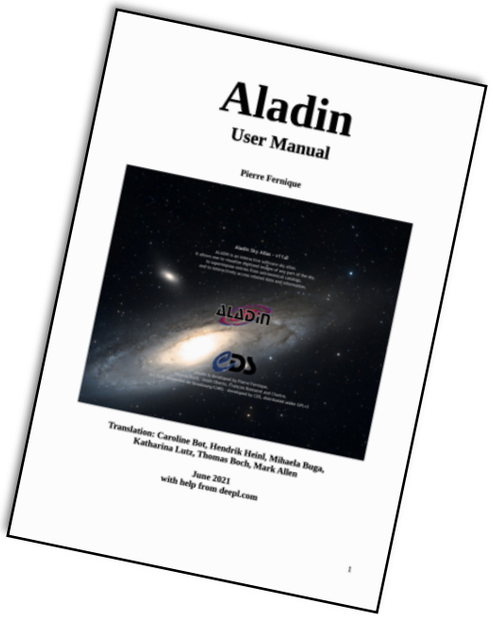
Special Session - Legacy and stewardship of astronomical archives towards multi-instrument, multi-wavelength and multi-messenger science.

Invited talk — ***Spatio-temporal exploitation of astronomical archives***, S. Derriere.

Other communication topics…

- Video hosting solution at UNISTRA
  - Will enable embedding

- Aladin Manuals in English (v10, 11)



- Best Practices for Data Publication in the Astronomical Literature:
  - ApJS submitted paper
  - Led by NED with many CDS contributions + others

Staff

Permanent staff:

11 - Researchers
(8 CNAP, 2 CNRS, 1 Emirat.)
9 - Software engineers
10 - Documentalists

Contract staff:

2 - Engineers
2 - Documentalists
2 - Postdoc researchers

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Direction and Administration

M. Allen (Director)

P. Fernique (Technical Lead), *C. Holtzinger*, **C. Halter**

Science

M. Allen
C. Bot
L. Cambrésy
S. Derriere
F. Genova *[Emirit]*
C. Loup
G. Monari
A. Nebot
P. Ocvirk
A. Siebert
B. Vollmer

Software Development and Operations

T. Boch
F. Bonnarel
P. Fernique
G. Landais
S. Lesteven
G. Mantelet
A. Oberto
F-X. Pineau
A. Schaaff

Documentalist

A. Eisele
M. Brouty
C. Brunet
M. Buga
E. Collas
M. Neuville
E. Perret
E. Son
P. Vannier
P. Vonflie
C. Fix
K. Van der Woerd

Support (shared with Observatoire de Strasbourg)

S. Langenbacher, V. Trimbour, C. Saillard, T. Keller, *M. Misslin*

Post-doctoral Researchers
S. Amodeo
J. Kuuttila

Project support
H. Heinl
C. Holtzinger

Ph.D Students
T. Lizee

Software Development Interns
7 interns, short term contracts

Staff - recent changes

Permanent staff:

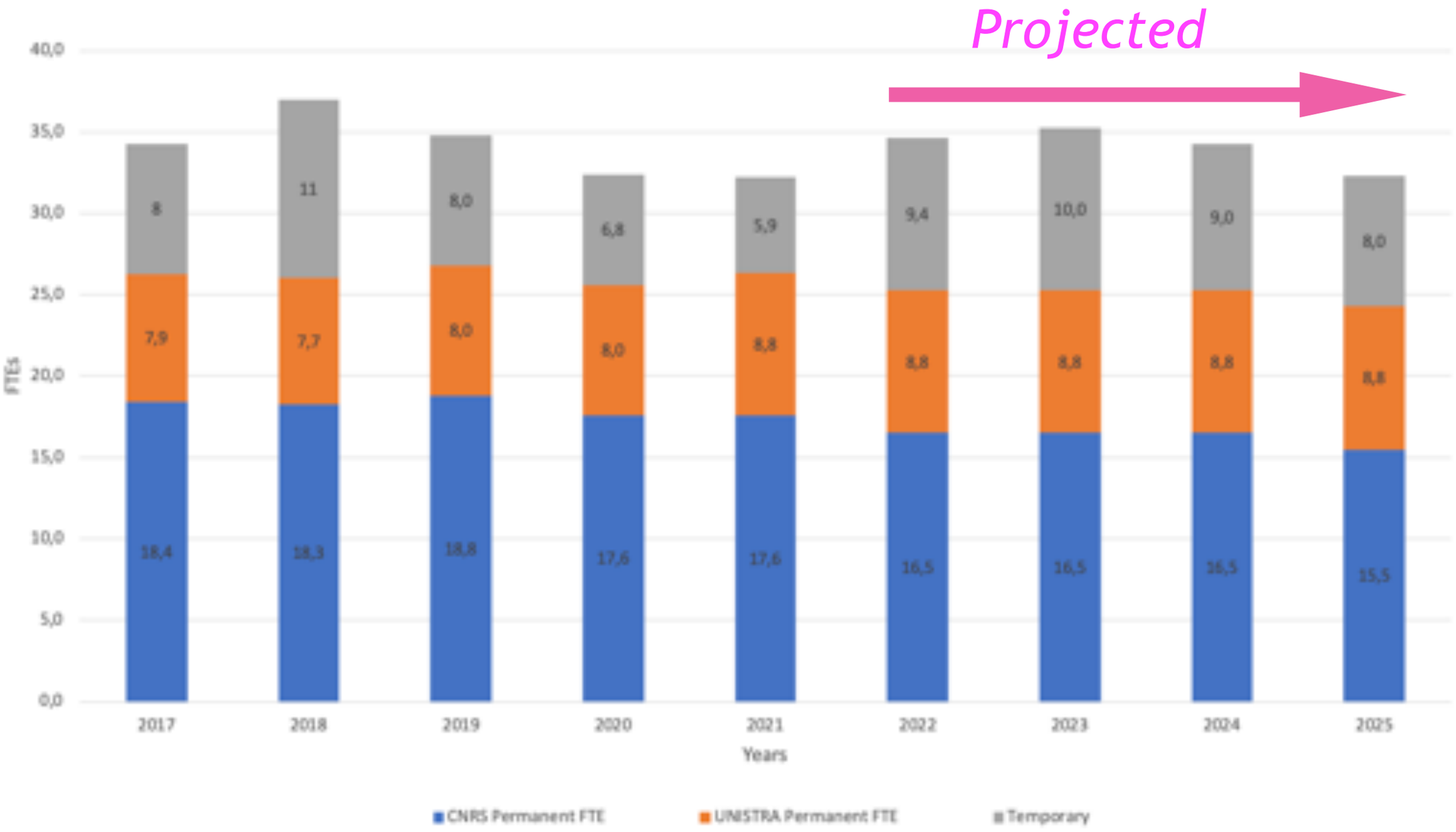
— CNRS competition for documentalist: **E. Collas** started November 2020

— Transition to Emeritus Status: **F. Genova** - December 2020
— Documentalist - **F. Marquis** - departed October 2021 *(25 yrs @ CDS)*
— Adminstrative assistant - **L. Arbousse** - departed April 2021 *(17 yrs @ CDS)*

Continuing UNISTRA contract: G. Mantelet CDI - started Jan 2021

Contract staff changes:
— ESCAPE engineer: **M. Baumann** - departed April 2021
— Postdoc **Y. Stein** - departed February 2021
— Postdoc **K. Lutz** - departed March 2021
— Postdoc ESCAPE science support - **S. Amodeo** - started March 2021
— Postdoc XMM2ATHENA - **J. Kuuttilia** - started August 2021
— IT Support Engineer — **M. Misslin** - started September 2021
— CDS Administrative Project Manager - **C. Holtzinger** - started October 2021
— Seeking a contract engineer for VizieR — now re-advertised

Full Time Equivalent



$$17.6 + 8.8 + 5.9 = 32.3$$

National and European Landscape

Elements that define high level policies:

- **French National Roadmap for Research Infrastructures**
- MESRI National Plan for Open Science - 2nd plan released in 2021
- European Cloud Initiative & European Open Science Cloud (EOSC) **
- (ESFRI Roadmap)

Recent and current developments

- CNRS-INSU Prospective
- INSU Astronomy & Astrophysics Prospective
- CNES - French Space Agency prospective
- US Decadal review — just released in November 2021
- ASTRONET Science Vision and Infrastructure Roadmap (in preparation now)

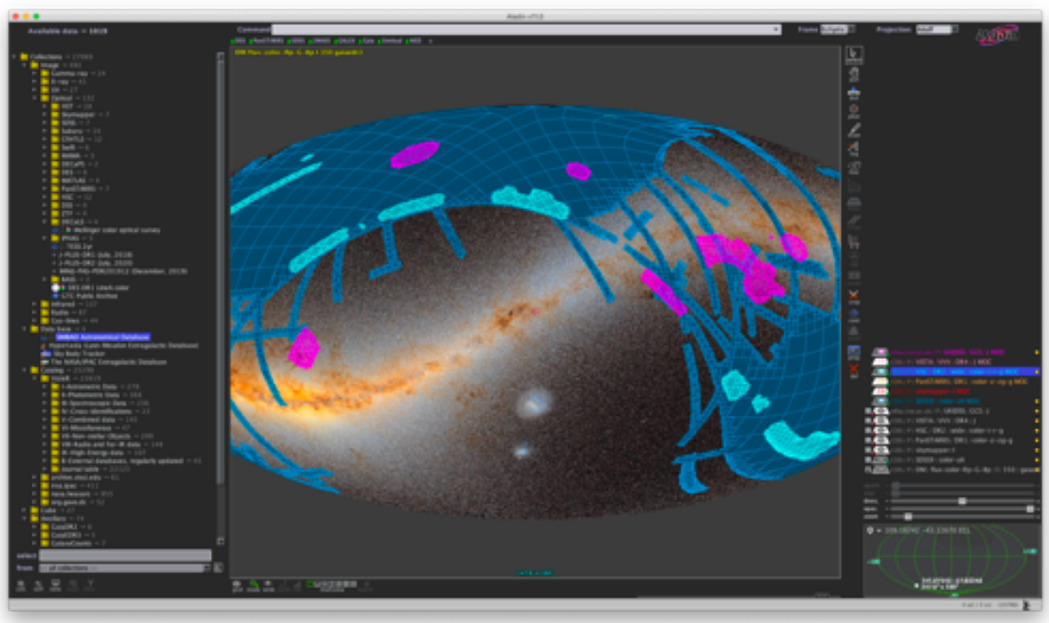
French national roadmap for Research Infrastructures

Successful application for inclusion

Criteria:

- European & International dimensions
- Scientific and Technological concept
- “Data” and “Data Management”
- Scientific production and service use
- Relationships to other infrastructures
- Industrial relations, innovation and societal impact

Previous version 2018



Application: Feb 2021, Result October 2021

Elements concerning scientific production and training

- Guidelines for users to cite the use of the services?
  - CDS service acknowledgements
- Number of publications from projects that have used the infrastructure (refereed)
    - 2017 : **947**
    - 2018 : **1107**
    - 2019 : **834**
    - 2020 : **1063**

  (ADS full-text searches on SIMBAD, Vizier, Aladin and manual validation)
- Number of communications: **40-68** per year in 2017-20
- Events for training/dissemination: **13-32** days per year in 2017-20
- Trainees at CDS: **465-710** person-days per year in 2017-20
- Descriptive text explanations: CDS processes **~12000** articles per year, and ingests from **1200-1500** catalogues per year

Elements about usage and users, requests for *'Estimates'*

- Number of external academic users: national and international
  - **2464, 32700*** *(2020)*
- Number of unique visitors of web interfaces: **44000** *(2020)*
- Number of accesses via APIs: **660 000 000** (2020)
- Service availability: **> 99%**
- * Complementary remarks: *Uncertainties explained*

Relations to industry, innovation and socio-economic impact

- Contracts that *'valorize'* the CDS - to enable/support community use of the CDS services (including observatories and space agencies).
- Contracts for publication of data and provision of information.
- Growing use of CDS services by planetaria, and developers of planetarium software *(e.g. Digistar)*.
- Training of interns - excellent preparation for engineering students.

Large project engagements

Euclid mission

- Presentation made to Euclid Consortium Board
- CDS individuals applying for EC membership
- MOU to be explored for CDS-EC-ESA

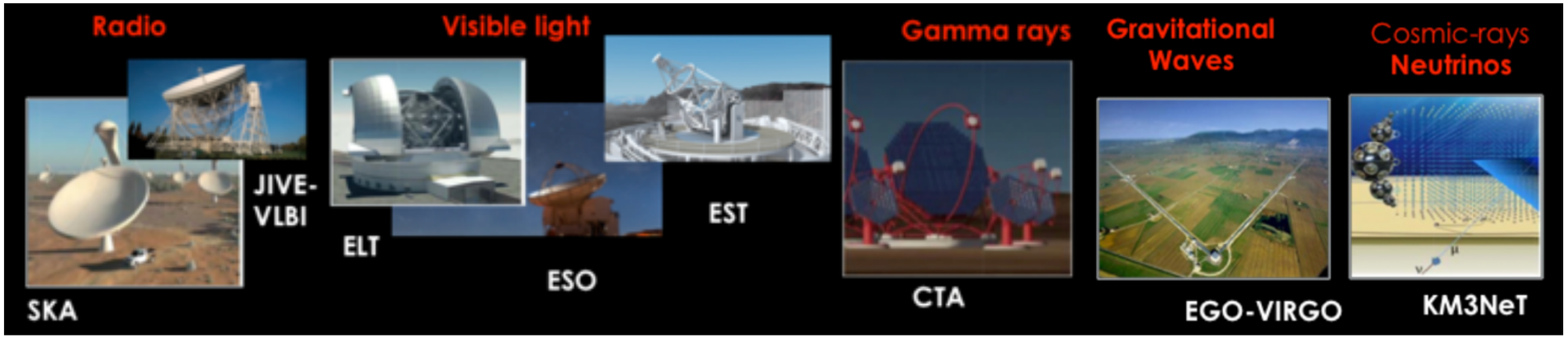
LSST Brokers

- CDS X-Match is part of the proposed FINK brokers
- Collaborative discussions ongoing with FINK and other brokers about thier use of CDS services as part of their operations

ESO and Hawaii observatoires

- Responding to requests for VizieR implementations as part of their operations

ESCAPE connections to ESFRI and other RIs



Telescope/instrument	Number of records	ETA	manageable at CDS? (tech-wise)	Band/depth
Gaia all epochs	10^9x100epochs~1e11	2022-2024	no	
Gaia flat	1.7x10^9	available in VizieR	yes	
Pan-STARRS DR1-2 all epochs	2.10^9x10epochs~2^10	done	~yes?	opti
Pan-STARRS DR1-2 flat	2.10^9	DR1 in VizieR	yes	optical wide
ESO Phase III cats	1-2x10^9	ongoing	yes	optical/NIR wide
LSST/Rubin all epochs	4^10x(~250epochs) = 10^13	2025-2030	no	optical wide
LSST flat	40x10^9	2025-2030	yes	optical wide
EUCLID	~10^10 galaxies	2028 and beyond	yes	optical/NIR wide
JWST	deep extragal surveys, narrow FoV, high res	2022 and beyond	yes	NIR deep
Nancy Grace RST	a few 10^9 stars, 4x10^8 galaxies	launch 2027	yes	orange / NIR deep/medium wide
ELT	~narrow FoV, very high res (0.005”)	2027	yes	optical/NIR deep
Giant Magellan Telescope	deep extragal surveys, narrow FoV, very high res (0.01”)	2029	yes	optical/NIR deep
TMT	deep extragal surveys, narrow FoV, very high res (0.015”)	2027	yes	NUV/optical/mid-IR
21cm / SKA	up to 10^9 gals	2030-ish?	yes	radio wide deep
GW: LIGO/PTA/LISA	?	operating and upcoming	yes	grav. wave
CTAO	?	2024	yes	gamma
KM3NET	?		yes	neutrino
ATHENA	several 10^5 - 10^6		yes	X-ray
SVOM	?		yes	gamma

Future-watch
- missions/surveys

Virtual Observatory and Projects

- **CDS participation in Virtual Observatory activities at National, European and International levels**
  - OV-France (&ASOV), Euro-VO
- **CDS continues leading role in IVOA**
  - Executive board, WG/IG roles
  - Recent changes:
    - **A. Nebot**: Chair of Committee for Science Priorities (CSP)
    - **G. Mantelet**: V. Chair Data Access Layer WG
    - **H. Heinl**: Chair Education IG

- **Important progress for the interoperability of astronomy data and services**
  - e.g. fundamental underlying indexing of space-time coverage of data and HiPS system for hierarchical approach to big data. CDS MOCServer 2
  - Leadership for integration of Radio Astronomy (+others) into VO framework
  - CDS Publishing registry and many more…

Projects

- **ESCAPE** - big project in progress… see next slides…

- **Europlanet 2024** Research Infrastructure (EPN-2024-RI)
  - Small but important CDS participation (~20PM)

- **XMM2ATHENA** **(***started 2021)*
  - A. Nebot is the ObAS coordinator. Postdoc — J. Kuuttila
  - ~5PM participation of CDS related to X-Matching

- **EOSC Future** - large (~40 M€) project of science and e-Infrastructures
  - Started April 2021
  - CDS is small part (20 PM) for Test Science Cases and training activities

ESCAPE Project

- **E**uropean **S**cience **C**luster of **A**stronomy and **P**article physics **E**SFRI infrastructures. (H2020 project, ~16 M€, 31 partners)
- Addressing the Open Science challenges of large astronomy infrastructures - in the context of EOSC
- CDS leads the CEVO Work Package (WP4)
  - **C**onnecting **E**SFRI to EOSC using **V**irtual **O**bservatory
  - Coordination of 16 partners (~348 PM, ~68 PM at CDS)
- Project extended to 48 months - ending Jan 2023
- Deliverables, milestones public on ESCAPE pages
- CEVO details on wiki pages
- Progress meeting/GA held Sept 28-29, 2021
- New project officer insists on more reviews - Next one 02 March 2022

https://projectescape.eu

Work Program

Data Lake:
Build a scalable, federated, data infrastructure as the basis of open science for the ESFRI projects within ESCAPE.

Science Platforms:
Flexible science platforms to enable the open data analysis tailored by and for each facility as well as a global one for transversal workflows.

Citizen Science:
Open gateway for citizen science o
ESCAPE data archives and ESFRI community

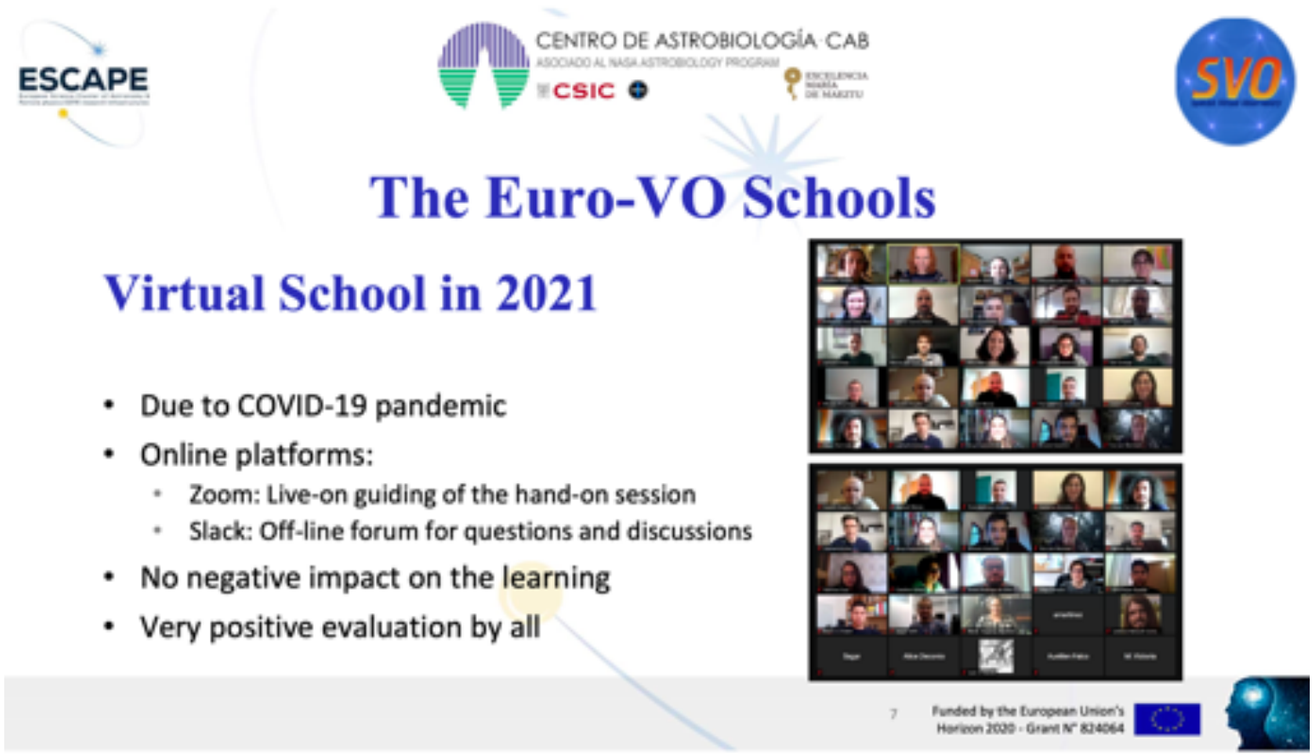
Software Repository:
Repository of "scientific software" as a major component of the “data” to be curated in EOSC.

Virtual Observatory:
Extend the VO *FAIR* standards, methods and to a broader scientific context; prepare the VO to interface the large data volumes of next facilities

Interoperability Standards
Metadata / Protocols
International context -

Funded by the European Union’s Horizon 2020 - Grant N° 824064

- **Virtual School Feb 2021**
  - Focus on students & postdocs
  - CDS tutorials, + others
  - Support of science cases



- **Technology Forum April 2021**
  - **Sharing practical expertise**
  - **‘Hack-a-thon’ mode**

- **Many project meetings… cross-WP activities, EOSC events…**
- **Some of the CDS benefits include:**
  - Aladin Lite v3, mocpy, MOC 2.0 reference implementations, +…
  - Support of training events - scientific and technical
  - Postdocs & engineers
  - High visibility in Open Science communities, and for European projects

WP4 Next Steps … cont.

* **2nd School** – Planned for Strasbourg, February 2022. (D4.6)
  * In-person if possible, otherwise virtual again
  * **Idea –** link the school with TSPs so that VO capabilities are well understood by TSP teams, plus precious feedback on the developments / tools / services

* **Final Deliverable Reports to be developed in 2022**
  * D4.7 Final analysis report on integration of VO data and services into EOSC
  * D4.8 Final analysis report on IVOA standards and stewardship best practices

* **Oct 2021 – June 2022** – analysis of the progress and results and prepare reports and publications, make visible in Astronomy and EOSC communities

Funded by the European Union's Horizon 2020 - Grant N° 824064

Supporting Open Science infrastructure and organisation in astronomy (and more…)

International Astronomical Union (IAU)

- **C. Loup** elected to Organising Committee of Commission B2 *“Data and Documentation”*
- CDS operates the IAU Dictionary of Nomenclature

International Science Council (ISC).

- Committee on Data (CODATA)
  - Task Group on Digital Representation of Units of Measure (DRUM) - input from **S. Derriere** bringing in experience of IVOA VOUnits standard

Research Data Alliance (RDA, RDA-France)

EOSC

- CNRS and UNISTRA are members of the newly created EOSC Association.
- **M. Allen** a coordinator of the Task Force for *Researcher Engagement and Adoption.*

Summary

- A challenging year for CDS
- Operational stability and high level production
- Important progress on core work of operating, maintaining and developing CDS reference services
- Adaptation of the ways of working
- Strong interactions and visibility in the community
- Progress on projects and new projects proposed
- Getting ready for a new era of Open Science
- Getting ready for the 50th anniversary of CDS 1972-2022

The CDS Information System

Overview & stats 2021

CDS Council – 29 & 30 November 2021

Pierre Fernique
on behalf of all the CDS staff

Information System update principles

The effort to **update the CDS's information system** continued in the direction initiated three years ago:

1. **Rationalisation** of the system's components
2. **Removal** or **evolution** of obsolete components
3. **Full redundancy** in distinct geographical locations (as far as possible)

This action must be driven by a precise knowledge of the functioning of the IS based on:

1. Operational **supervision**
2. **Monitoring** of use and content

CDS I.S. = 3 main services

Objects
from litterature

SIMBAD

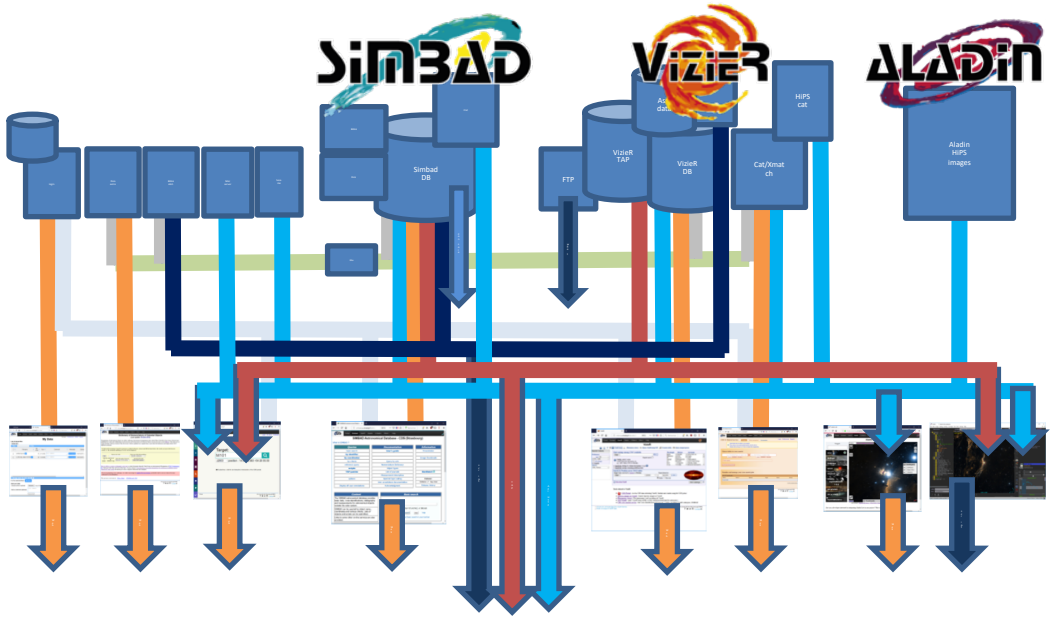
Catalogues
from litterature & surveys

VizieR

Images
from surveys

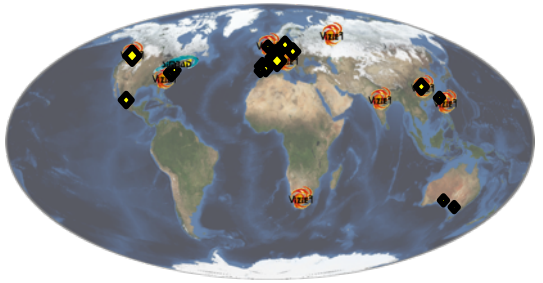
ALADIN

CDS infrastructure system 2021



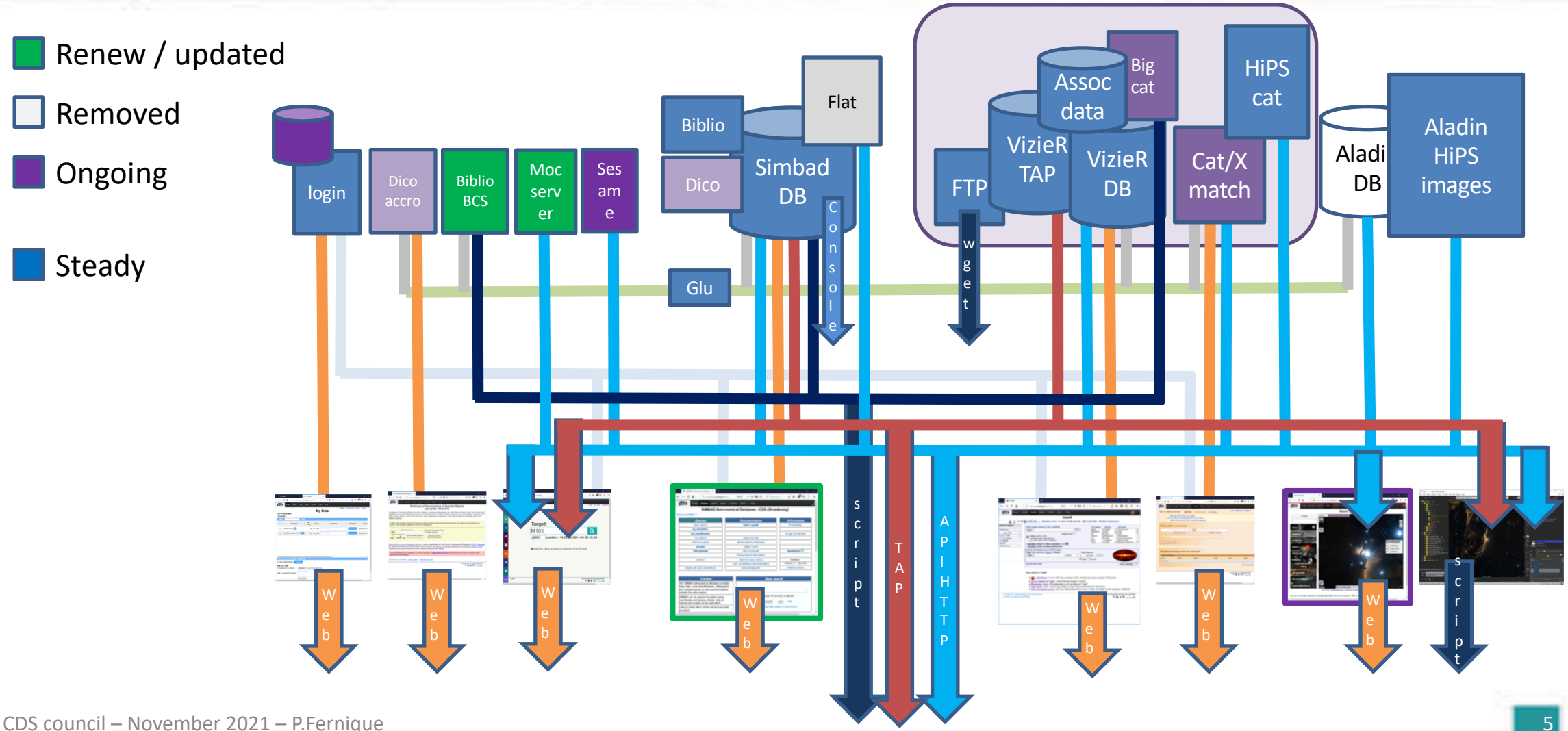
- **19 components** (DB, servers…)
- For various Web clients, APIs & clients

- **12 local machines** (real or virtual) for implementing the 19 sub-components
- CDS **allsky storage system**
  => 1.6PB replicated
- **2 computer rooms**
  => One at the Observatory
  => Replicated elements in Unistra Data Center
  - Partially done in 2021



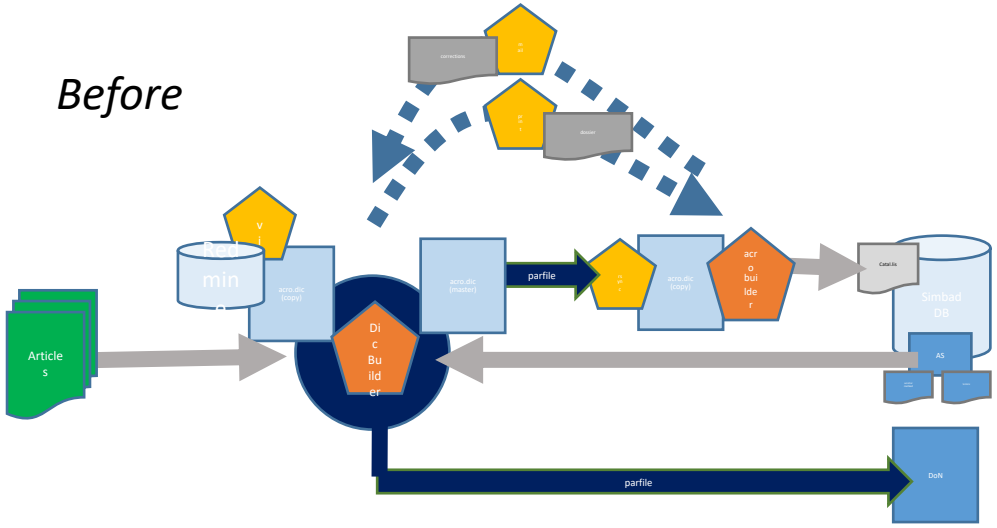
- **6 external sites** (5 VizieR + 1 Simbad mirrors)
- **20 partner HiPS sites** (Aladin tiles)

Evolution of IS components



Last example: Dictionary of Nomenclature redesign

Before



Good points

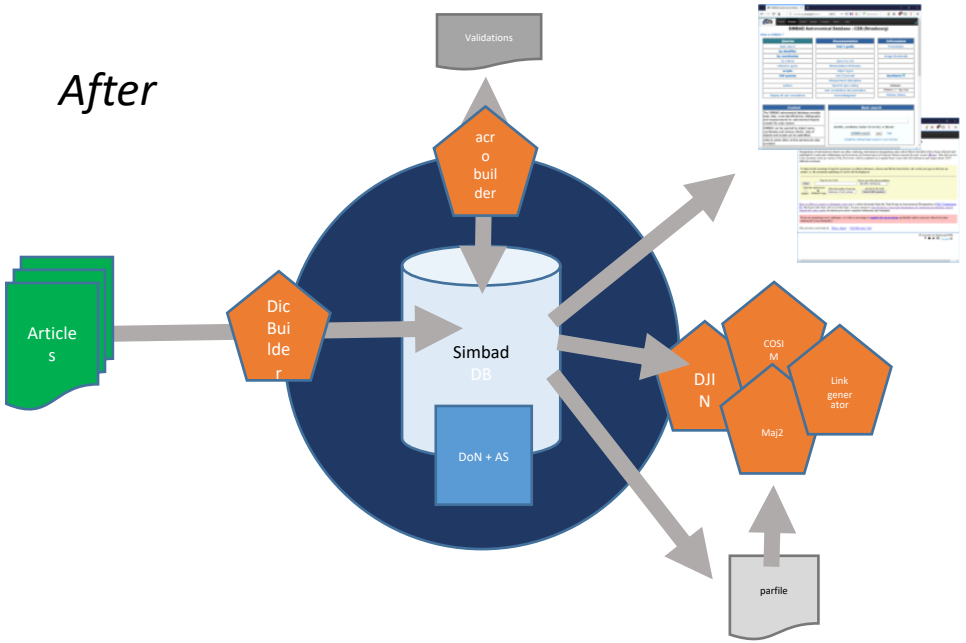
- A single file, very practical, very solid and durable
- A very precise description, perfectly adapted

Bad points:

- A tables flow and complex workflow for rather simple actions (=> update of 26000 ASCII entries)
- A parfile that is reaching its limits, mainly the shared access, but also some fields that are a bit too "house specific"
- Several output entities: Simbad Acronyms, Dictionary of Nomenclature, annex files=> multiplication of tools, round-trips, synchronization, cascading maj...

Description of the existing situation, Improvement analysis, Team discussion…

After



New design:

- Use the Simbad database as a master for the management of all acronyms (AS + DoN + link management), even if only one subpart is visible for Simbad acronyms (a flag in the record)
- All semantic content is kept (a priori)
- Syntax/structure BD to be studied

New architecture, Human resource, Road map (prototype, calendar…)

Information System recent evolutions

- Redesign of the **objects nomenclature service**
  => re-ingestion in the SIMBAD database (see previous slide – *in progress*)
- Deployment of **BCS** (CDS biblio)
  => for DJIN and VizieR (*nearly completed*)
- **MocServer 2** (*complete*) = CDS yellow page service
  => adding temporal coverages
- **Sesame 2** *(in progress)* = CDS name resolver
- **VizieR 2** : thinking/redesign of our catalog service (*long term action*)

- Stopping/evolving **old/redundant services** (*in progress*):
  => Aladin legacy image server, simplay, thumbnail image service

CDS inter-service developments

- Rewriting of **CDS web pages** (*in progress*)
  => Hugo/gitlab solution under evaluation)
- Replacement of the **CDS hotline** by a ticket tracking tool
  => GLPI evaluation (*in progress*)

- Rationalization of our **source code tracking** (*in progress*)
  - Gitlab CDS for all our codes
  - Github Microsoft for a few public lib/tools (Moc, Aladin Lite, ...)

CDS Web evolutions

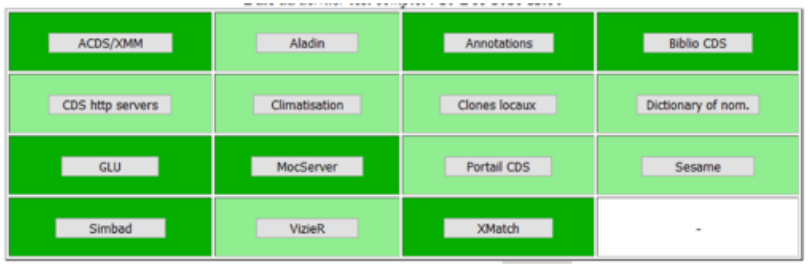
- **HTTPS** support *(in progress)*
  => required for external widgets (Aladin lite) + better Google indexing
- **Unique domain name** "cds.unistra.fr" (*in progress*)
  => better consistency, indexing and control of HTTPS certificates

Service Continuity - Disaster Recovery Plan

- CDS basic rule: **Total duplication of data** in at least 2 distinct geographical locations (*partial*):
  => PRA + load balancing...
  - **Unistra DataCenter** for an alternative site: alaskybis + 1 VM hypervisor
  - Regularly check that at least one mirror of VizieR and SIMBAD can be used to rebuild the service (in case of a major disaster at Obs) *(to improve)*
  - Think about duplicating the last unduplicated services: Xmatch + queryCat + CDSlogin + IRODS (*todo asap*)
- **Survival files**: 37 cards (*operational*)
  => simple and efficient method for minor failures

Operational management

- Close and efficient **synergy with the Obs infra team**
- **Supervision** (*fully operational*):
  - Global => GluSupervisor
  - Detailed => specific to each service (nagios...)
  - Technical (hardware) => zabbix



- Statistics and **activity measurement** tools (*fully operational*):
  - Adjustment of the Matomo tool configuration (GRPD compliance)
  - CDS dashboard

CDS dashboard

Statistics on all CDS services at a glance

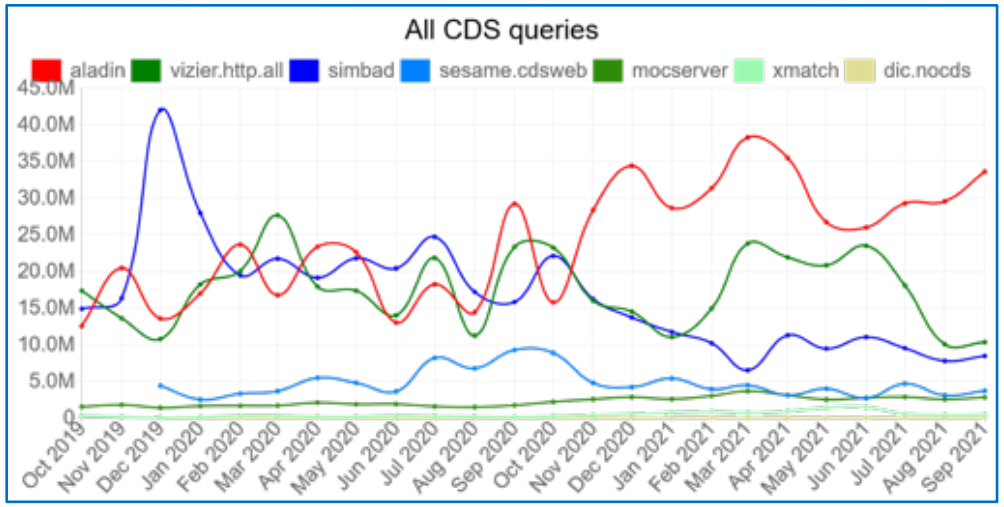
CDS main metrics *(Oct 2019 - Sep 2021)*

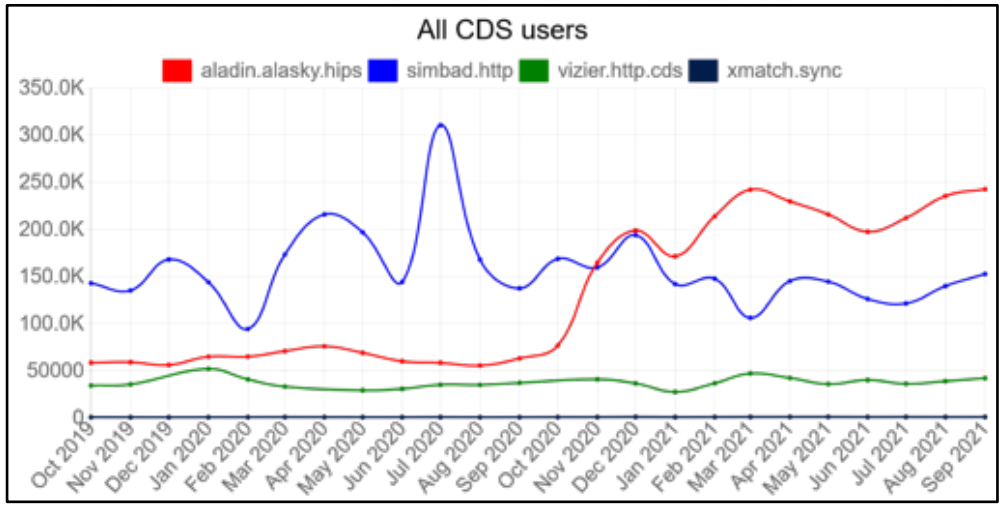
compared to Oct 2017 - Sep 2019

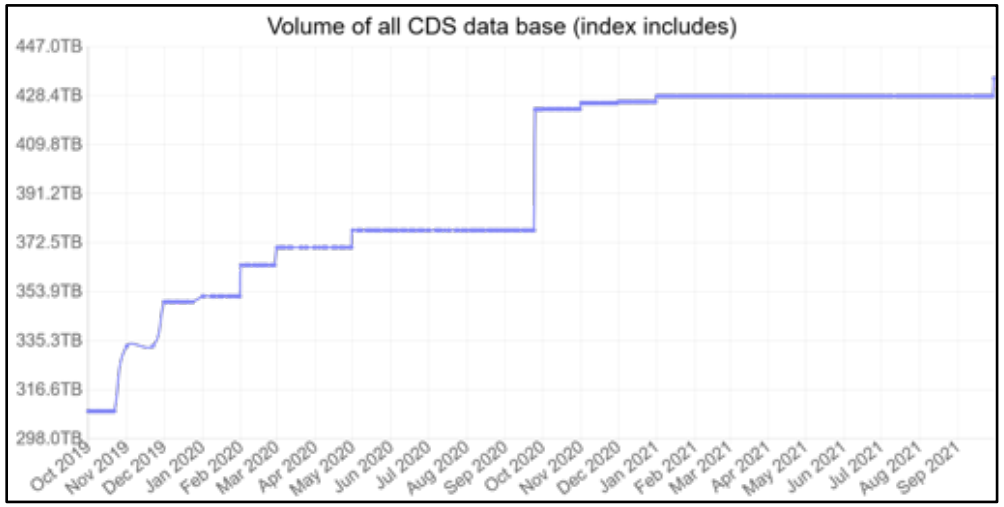
Main Services →	SIMBAD	VizieR	Aladin	Total (main services)
users / month	**146k** (-13%)	**30k** (-10%)	**230k**	**>230k**
queries / day	**393k** (-46%)	**569k** (+9%)	**925k**	**1.9M**
load / day	**5.3 GB** (-34%)		**100 GB**	**>105 GB**
data volume	**27 GB** (+12%)	**53 TB** (+14%)	**390 TB**	**435 TB**
data content	**12.6 M obj.** (+7%)	**21.3k cats** (+3%)	**990 HiPS**	
reliability	**99.44 %**	**99.65 %**	**99.61 %**	**> 99.44%**

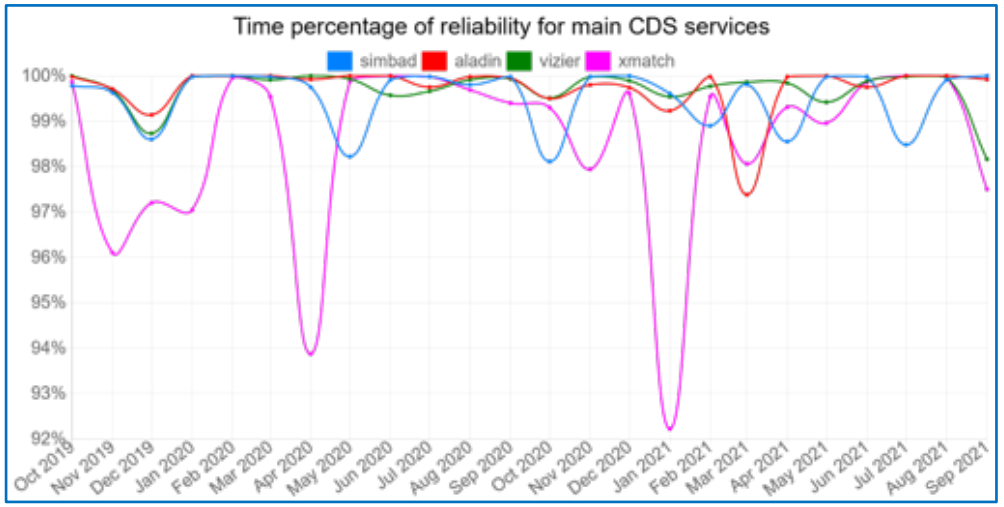
(Bots have been removed of these stats)
users/month Based on unique IP (Adv.Users = only based on "advanced scientifical queries") on the main site
queries/day Based on effective DB queries (object queries, catalog queries, HiPS tile queries...)
load/day Downloaded from CDS servers
data volume Volume of service (data + index)
data content Number of service "items"
reliability Percentage of service availability (mirrors included)

CDS main metrics 2 years evolution

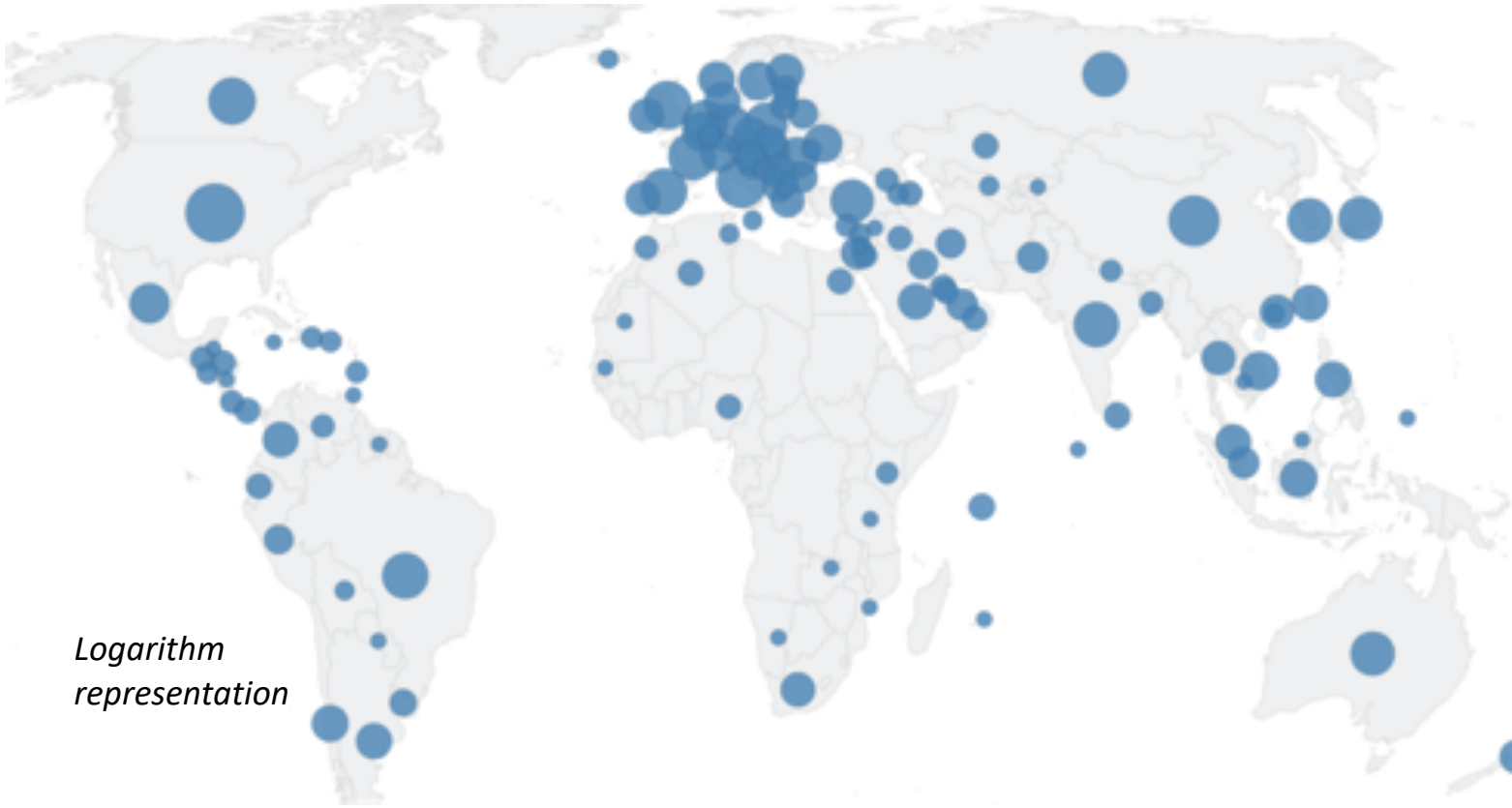




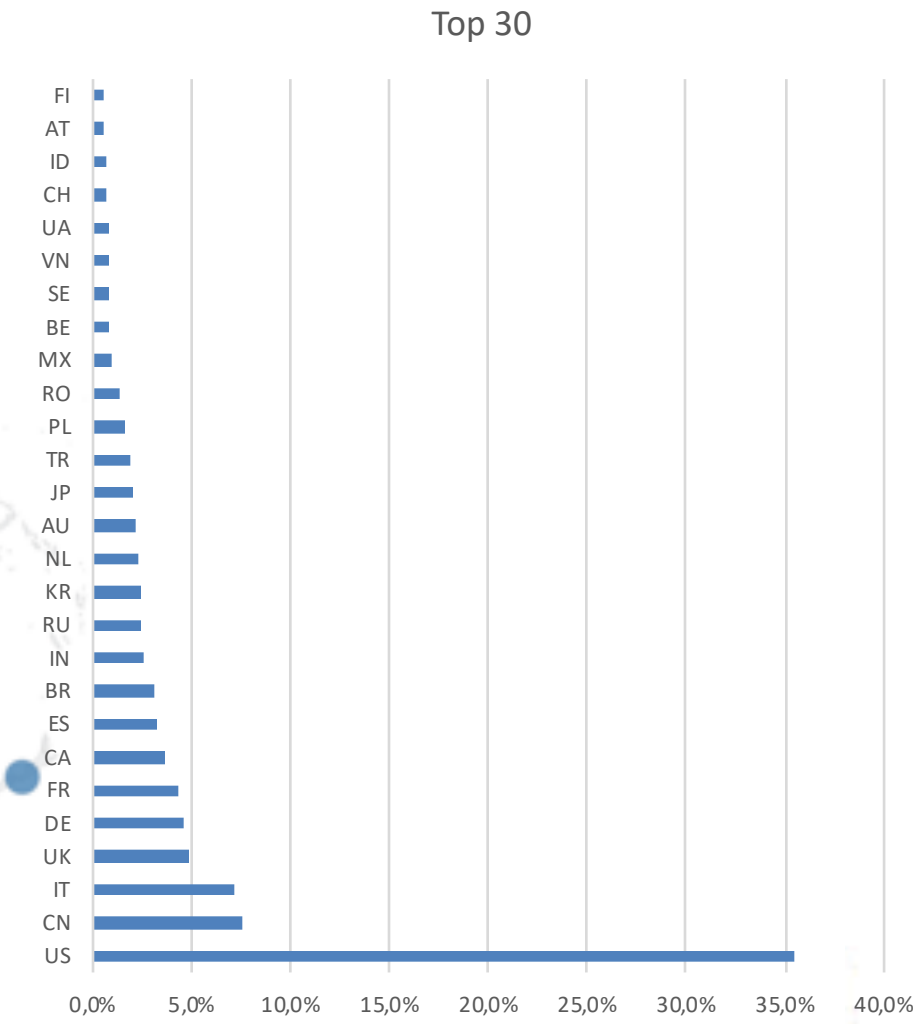




Country repartition (by unique IP)



Logarithm representation



Questions ?

The CDS Information System

SIMBAD : the bibliographic database

A meta-compilation of astronomical objects of interest that have been studied in the literature

CDS Council
November 2021

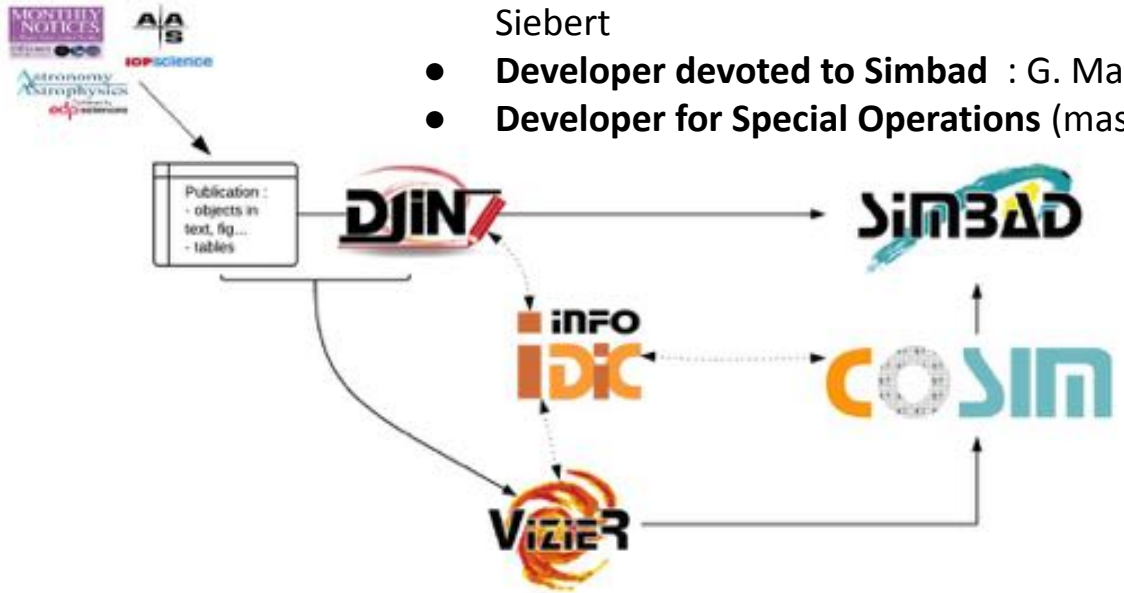
CDS team

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

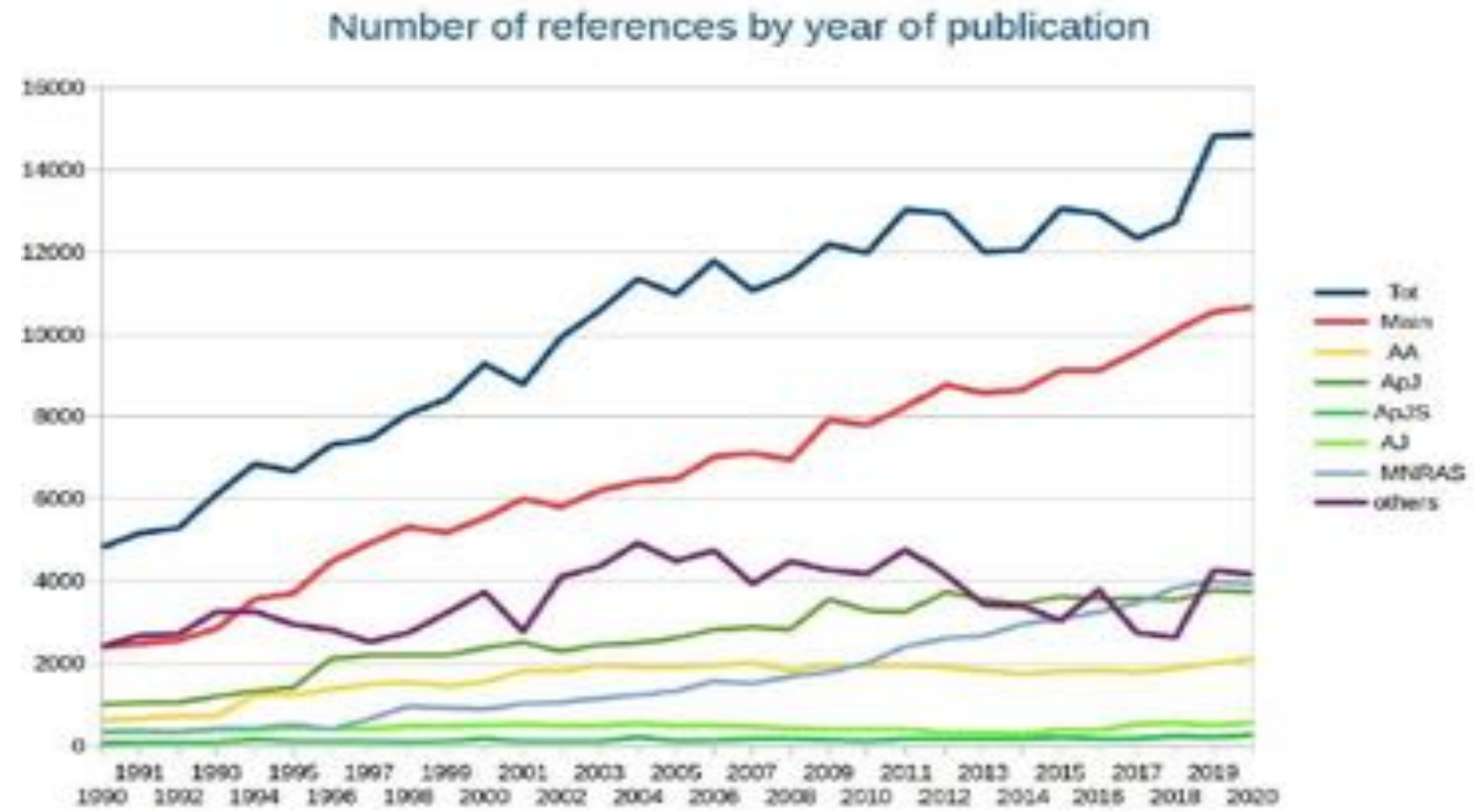
The Team

- **Bibliography & coordination** : S. Lesteven (engineer)
- **Scientific content** : C. Loup (astronomer)
- **Database & softwares** : A. Oberto (developer)
- **Nomenclature** : B. Vollmer (astronomer)

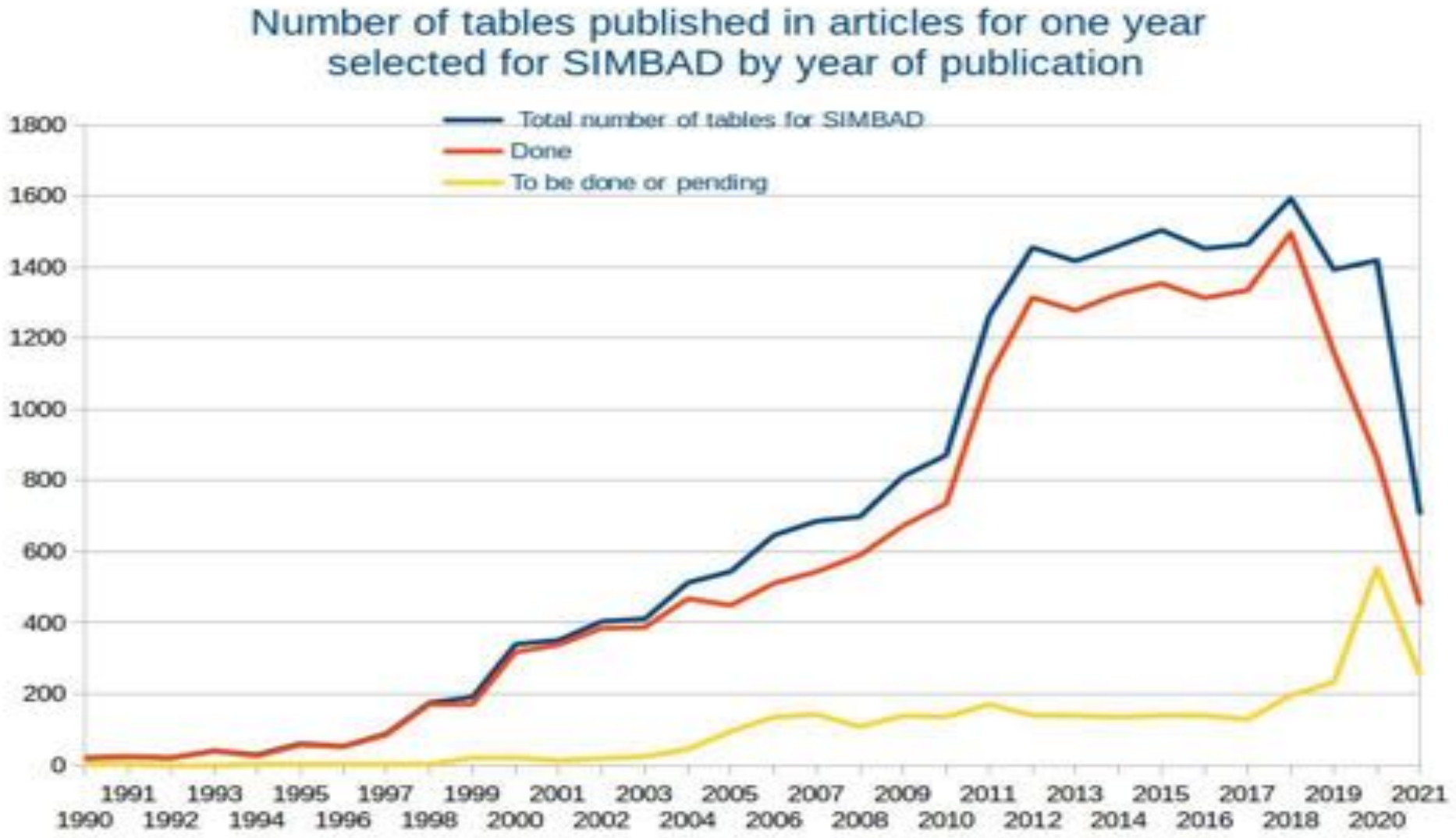
- **Documentalists** (Data Stewards) :
  - Nomenclature : M. Brouty, F. Marquis
  - Ingestion of references via DJIN : Aline Eisele, Evelyne Son, Magali Neuville, Philippe Vonflie
  - Ingestion of lists of objects via COSIM : C. Brunet, E. Collas, F. Marquis, K. van der Woerd, M. Buga, E. Perret
- **Astronomers** : C. Bot, L. Cambrésy, S. Derrière, F. Genova, G. Monari, A. Nebot, P. Ocvirk, A. Siebert
- **Developer devoted to Simbad** : G. Mantelet
- **Developer for Special Operations** (massive Xid) : T. Boch



Workflow : References



Workflow : objects



The Content in Numbers

- References : **15,000** (total 397,000)
- Objects : **1.3 million** (total 12.9 million)
- Identifiers : 8.2 million (total 45.1 million)

- Articles appraised : 730
- Added Priority 1 : 520
- Ingested : **735**

Increasing number of objects of interest in tables :

- From large astrometric, spectroscopic, variability, surveys
- 36 publications with more than 100,000 objects ingested in SIMBAD
- ⅓ from articles published from 2018 to 2021

Reorganisation of Object Types :

- Fully integrated in COSIM
- Automatic storage of objects with contradictory classifications in the literature

  --> very fast scientific expertise

  --> maintenance of the quality of the scientific content easier and at a high level

Special Operation on the TESS Input Catalogue (TIC) :

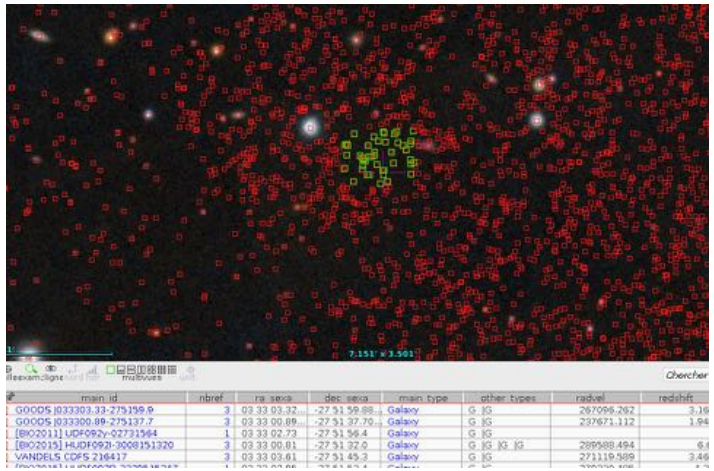
- New planets or candidates + expected large catalogues of variable stars
- 4 million TIC identifiers added on the basis of the 2MASS identifier to facilitate the ingestion of the new references and tables of objects

COSIM & Otypes

Data coming from a table

```
%I.0 [BIO2015] HUDF092B-3046852195
%I.R HUDF092B-3046852195
%J 03 33 04.68 -27 52 19.5 (NIR) C
2015ApJ...803...34B
%M.J (AB) 27.11 D 2015ApJ...803...34B
%V z:p  3.45 (NIR) E
2015ApJ...803...34B
%O.velocities |z|  3.45 | |p,,NIR,
|2015ApJ...803...34B
%C Galaxy
```

Test compatibility of all data with SIMBAD data



List of **compatibles types**

Object type code	Long code	Candidate code	Priority	Description	Comments	Hierarchy	Pragmatic compatibilities
G	Galaxy	G?	40	Galaxy		GGG	gLe, LeI, LeG, LeQ, NIR, Opt, GGG
LSB	LSB_G		10	Low Surface Brightness Galaxy		LSB	G, LSB, GiP, GiG, GiC, BiC, gLe, LeI, LeG, LeQ, NIR, Opt
bCG	BlueCompG		11	Blue Compact Galaxy		bCG	G, bCG, GiP, GiG, GiC, BiC, gLe, LeI, LeG, LeQ, NIR, Opt
SBG	StarburstG		12	Starburst Gala		SBG	G, SBG, H2G, EmG, GiP, GiG, GiC, BiC, gLe, LeI, LeG, LeQ, NIR, Opt
H2G	HII_G		14	HII Galaxy		H2G	G, SBG, H2G, EmG, GiP,

Bibliographical Center

- Evolution EDP Science on storage structure
- Evolution AAS Journals on download process
  - Missing tables from months
  - Access difficulties on IOP server
- New journals processed
  - MNRAS-Letters
  - PASJ
  - Science (in preparation)
- Same storage and conversion in XCDS for SIMBAD & VizieR

WEB

Evolution of web page to put ahead CDS services

Number of rows : 25

Want to see more from a catalogue? You can use VizieR to search in the same area for instance: Gaia DR2, 2MASS, AllWISE, SDSS, others

Show 100 entries

Search:

N°	Identifier	dist(asec)	Otype	ICRS (J2000) RA	ICRS (J2000) DEC	Mag
1	HD 86161	0.00	WR*	09 54 52.9034376427	-57 43 38.273623332	7.94
2	[DAB2013] feature 1	55.86	Rad	09 54 55.0	-57 42 45	
3	IRAS 09526-5729	167.56	FIR	09 54 32.2	-57 43 14	
4	IRAS 09529-5731	194.87	IR	09 54 36.2	-57 46 00	
5	IRAS 09535-5726	220.57	IR	09 55 13.2	-57 41 10	
6	TYC 8606-860-1	224.10	*	09 54 54.1545060518	-57 47 22.147682176	
7	TYC 8606-974-1	256.72	*	09 55 13.7863267390	-57 46 53.141794040	
8	IRAS 09525-5730	326.06	*	09 54 14.0722090879	-57 45 16.568777327	
9	TYC 8606-273-1	333.81	*	09 54 26.7704909782	-57 39 18.396309267	
10	TYC 8606-788-1	345.57	*	09 54 57.5932614299	-57 37 54.755730156	

HD 86161 -- Wolf-Rayet Star

Other object types:

ICRS coord. *(ep=J2000)* :

FK4 coord. *(ep=B1950 eq=1950)* :

Gal coord. *(ep=J2000)* :

Proper motions mas/yr :

Parallaxes (mas):

Spectral type:

Fluxes (B) :

SIMBAD Query around within arcmin

All (CDSPortal)

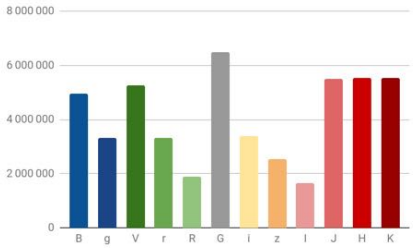
Send to

Photometry within arcsec

Dictionary brainstorming

- Status of dictionaries (IAU & SIMBAD definitions)
- Content to take into account:
  - Acronym
  - Format
  - Type of object
  - Synonyms
  - Tables
  - …
- Prototyping a new version

CONTENT UPDATES



IDENT
45 M
+23 %

FLUX
52 M
+12 %

BASIC
12.8 M
+4 %

ALLTYPES
25 M
+7 %

AUTHOR
132 K
+4 %

REF
396 K
+5 %

HIERARCHY
2.4M
+12 %

mesVar
2.2 M
+24 %

mesVelocities
8 M
+11 %

mesIUE
66 K

mesDistance
4.5 M
+13 %

mesRot
94 K
+75 %

mesHerschel
35 K

mesDiameter
6 K

mesSpT
1.1 M
+9 %

mesXMM
3 K

mesPM
6.8 M
+19 %

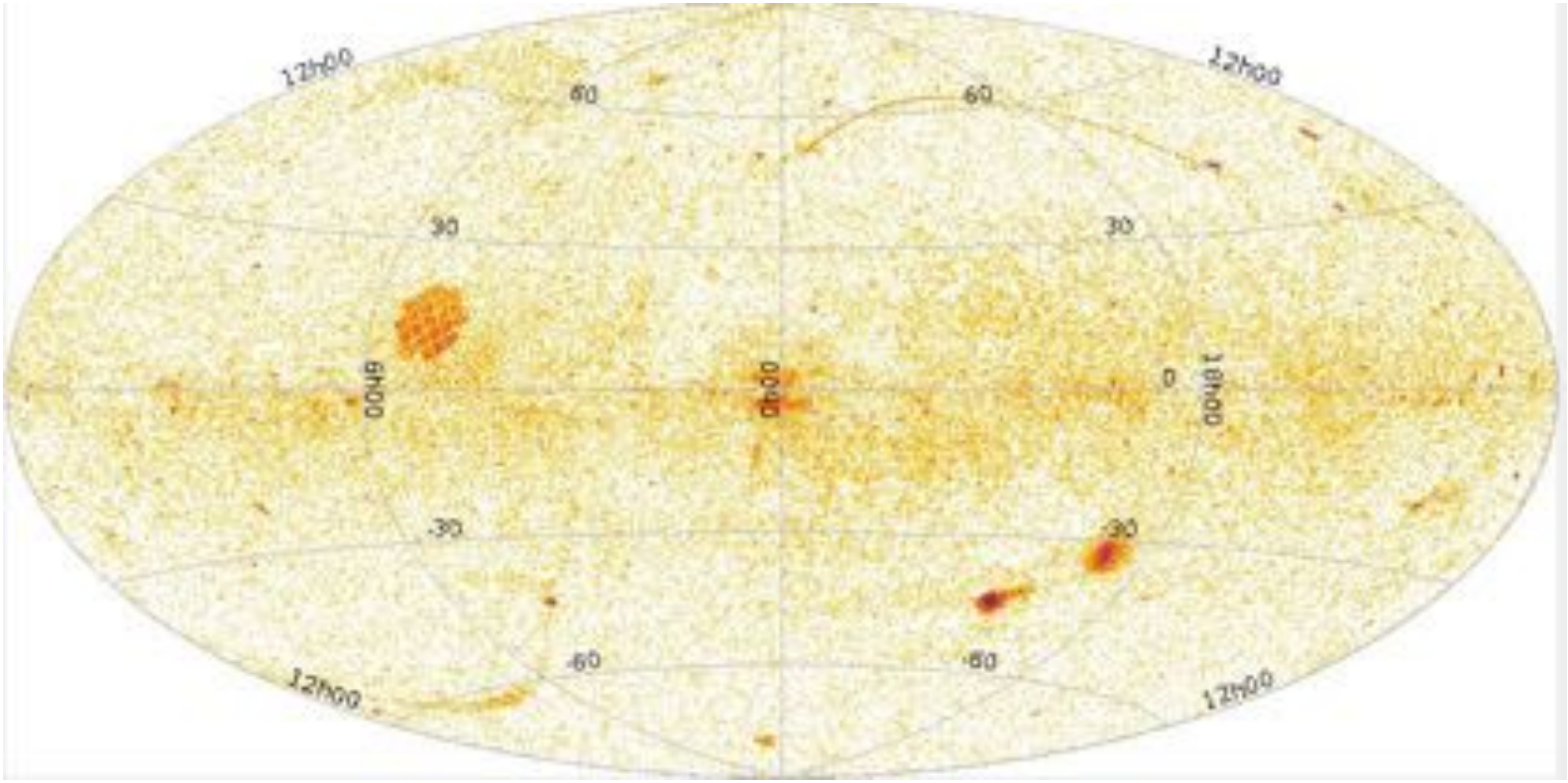
mesPlx
7.7 M
+16 %

mesISO
29 K

mesFe_H
1.4 M
+43 %

CONTENT UPDATES

Since 2020-oct, location of objects updated



Scientific council 2021

VizieR Staff and contributors:

Astronomers: P.Ocvirk, C. Bot, G. Monari, S.Derriere
Engineers: G.Landais, F.X.Pineau,T.Boch
Documentalists: P.Vannier, E. Perret, C. Fix, M. Brouty

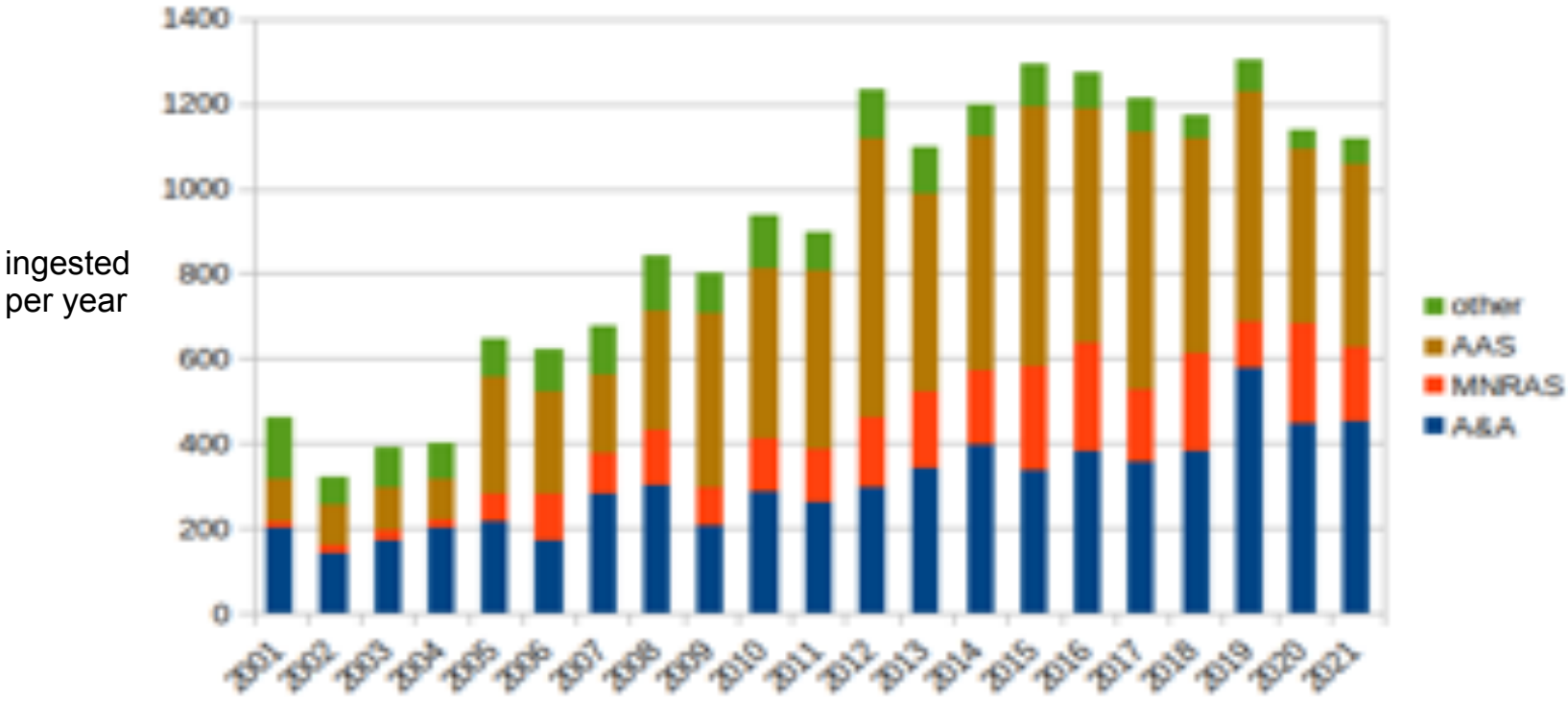
Non-CDS: L. Michel, C. Saillard, T. Keller (Strasbourg Observatory)

VizieR content - I

VizieR content - II

Ingestion statistitics :

- A steadily increasing volumetry : +1123 in past year
  - Total : 21,421 catalogues, 48,164 tables, (51 billion records)



VizieR content – III : Very large catalogs & co

Very large catalogs ingested in past year:

- Gaia EDR3
- GPS1+
- ESO VHS DR5
- ESO VEXAS DR2
- ATLAS-REFCAT2
- Gaia EDR3 distances (long)
- Gaia DR2 Extinctions (long)
- Re-ingestion of CatWISE following correction of positions

"Thick" catalogs: > 150 columns

- RAVE DR6 (several 100 columns)
- RELICS (150 cols from 46 tables)

In progress:

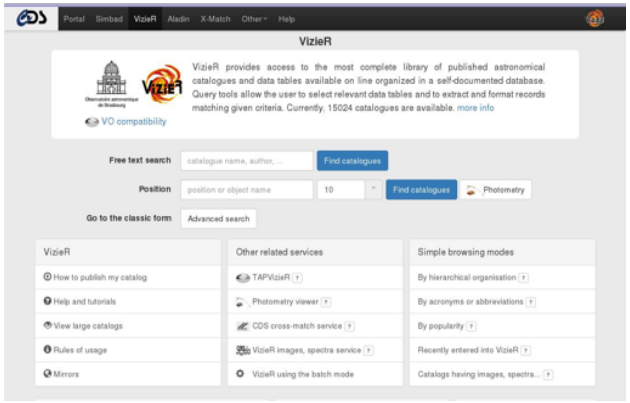
- SDSS DR16 (final stages)
- Pan-STARRS DR2 (contact re-established)

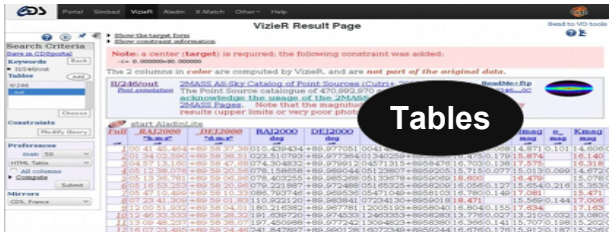
Planned for 2022:

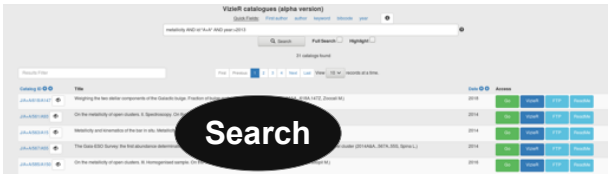
- Gaia DR3
- GSC 2.4 (3.5 billion objects)
- ESO phase 3, new DRs

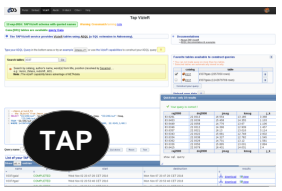
Accessing VizieR data

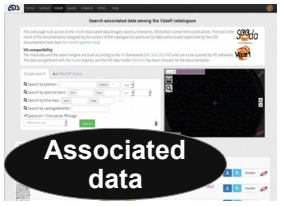
VizieR services and access modes: discovery, table access, visualization, photometry, associated data, astroquery, ...

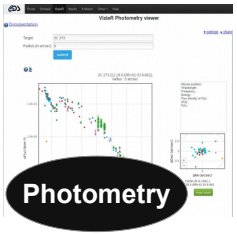


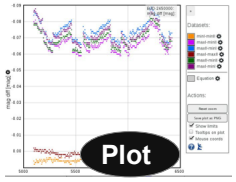




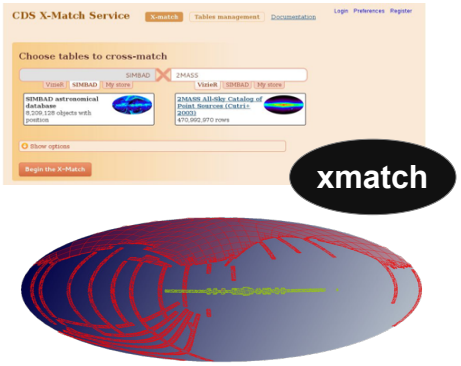


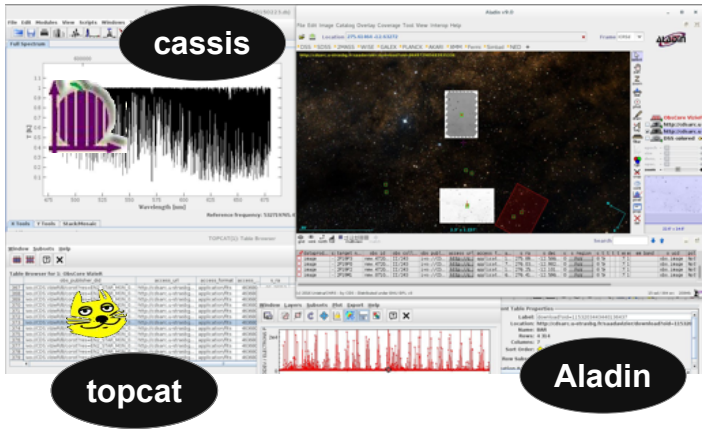






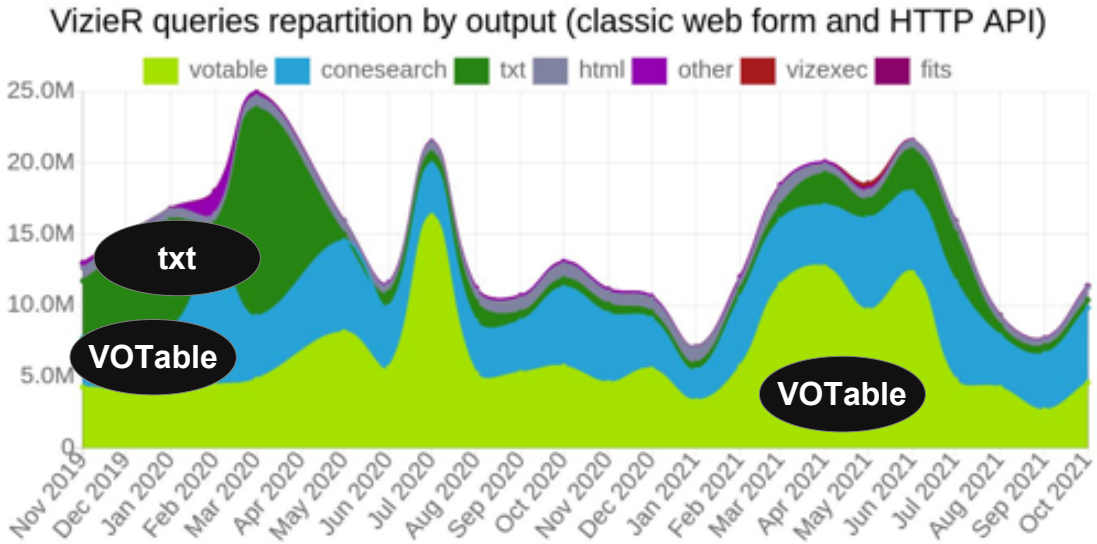
Derived products provided in CDS or data available through software, API..

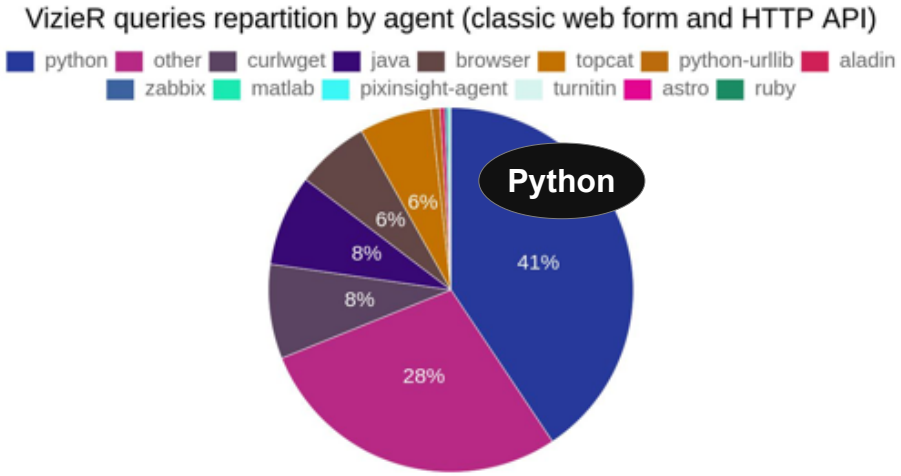




VizieR usage statistics

VizieR statistics (from the CDS statistics collector)





Total : ~533K queries/day

TAP : ~16.7K queries/day

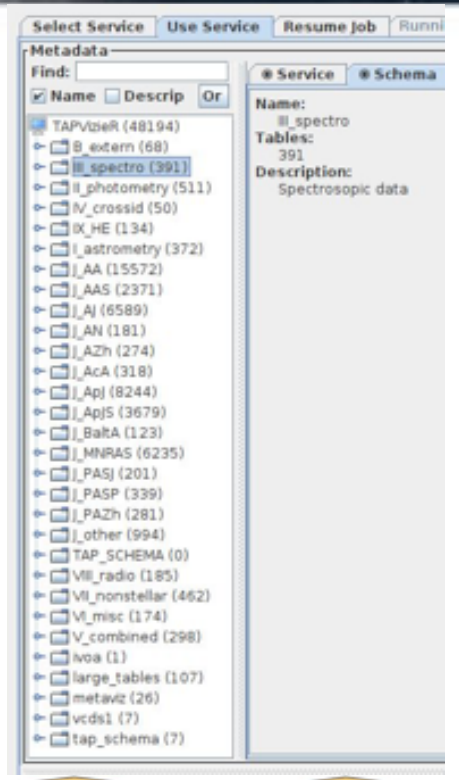
Associated data (spectra/images): ~1,200 queries/day

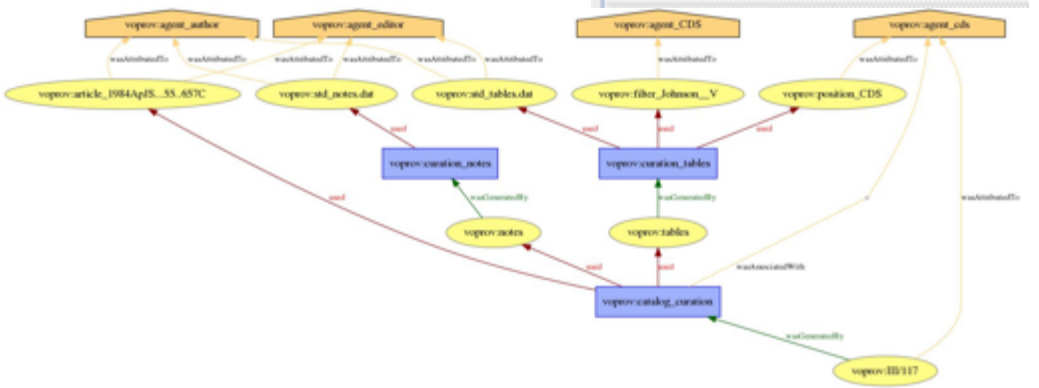
Importance of the Virtual Observatory (conesearch+VOTable)~70% of the queries

Importance of Python(41%), TOPcat (6%), curl+wget (10%)

Recent activity

- Solar System catalogue built with VOParis (VESPA): collection of VizieR catalogue gatered into a single table , curated by Paris
- The TAPVizieR categories have been re-organized following the VIzieR category
- The VO registry entry for VizieR has been updated and now includes the catalogues' MOCs, spectral range, mirrors and DOIs.
- Cdspyreadme library improvment for authors to generate CDS and MRT table format (AAS) - available in guthub, pip (PyPI)
- Workflow update:
  - workflow process improvment to fill the basics catalogue metadata (FX.Pineau) : UCD, description, units ..
  - Adapt the article xml extraction (getapj) to use the BCS (Bibliographic Center Supervisor, G.Mantelet)
- Provenance catalogue information available in landing page
- Very first steps of the vizier ingestion code renewal :
  a long work just beginning to put the vizier source code in the long term -
  (difficult to do by a unique person due to the heavy maintenance)





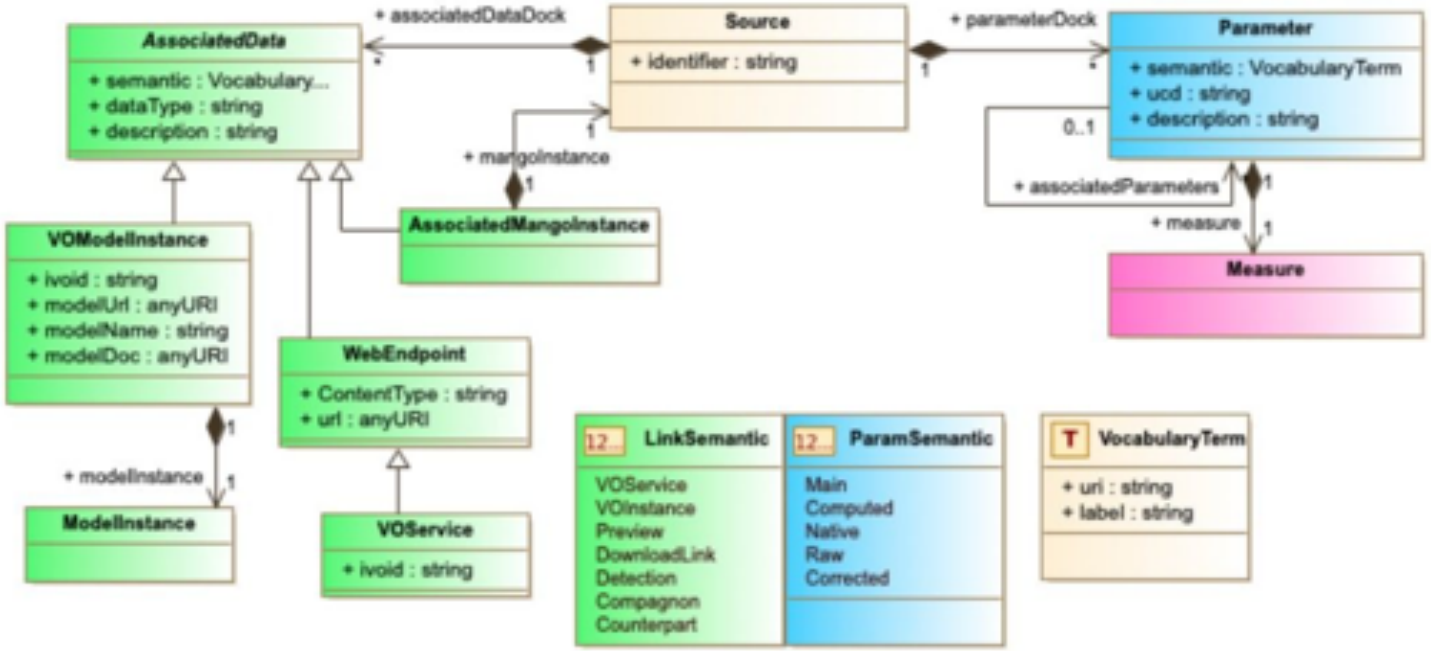
Recent activity

Other works :

- Exploring rich VO output using Data-models including filter and Provenance information (L.Michel, F.Bonnarel, M.Louys)

  build an output (VOTable) to higlight measured (eg: photometry) with precise metadata

- Exploring large table access in PostgreSQL



Outlook – 2021 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, new engineer required)
  - Homogenization of pipelines (large catalogs, journal catalogs)
- Very large datasets: High Cadence Wide Field surveys (e.g. LSST)
  - LSST: 40 billion objects x 250 visits = 10^13 records
  - Gaia DR3-4 /PS DR2: 2x10^9 objects 100 visits = 2x10^11 records
  - Should CDS develop such capacity?
  - How commonplace / rare will such datasets be in 10 years?

Aladin:

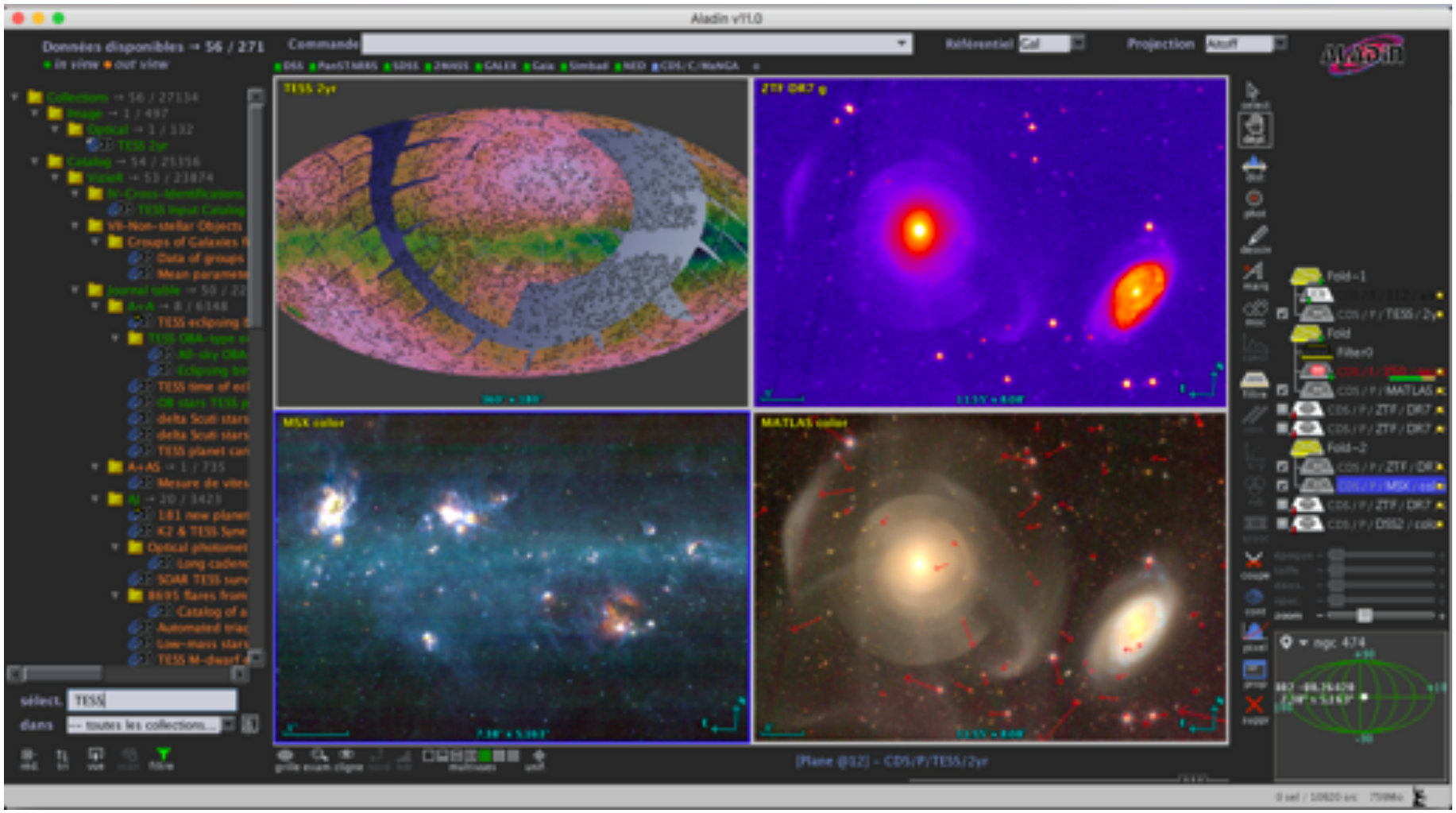
Highlights, statistics, perspectives

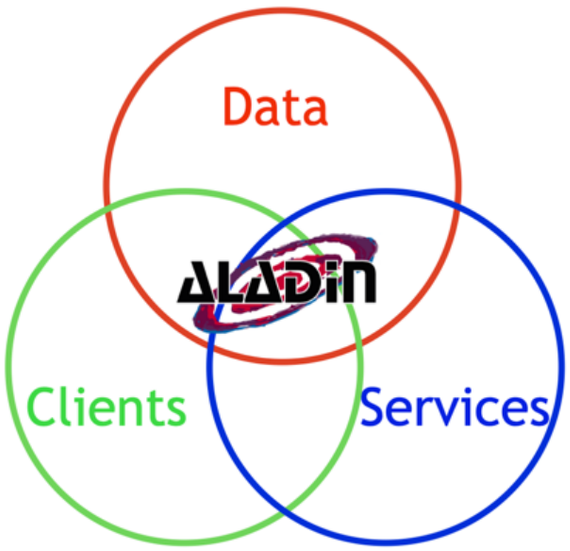
CDS council - November 29, 2021

Thomas Boch, Caroline Bot, Pierre Fernique, Mark Allen, Matthieu Baumann, François Bonnarel Mihaela Buga, Sébastien Derriere, Hendrik Heinl, Katharina Lutz, Ada Nebot, François-Xavier Pineau, Christophe Saillard, Thomas Keller

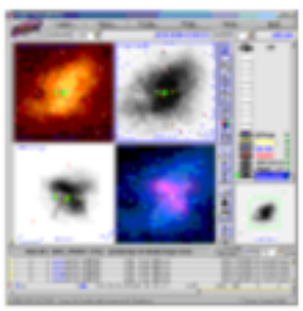
CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Aladin project





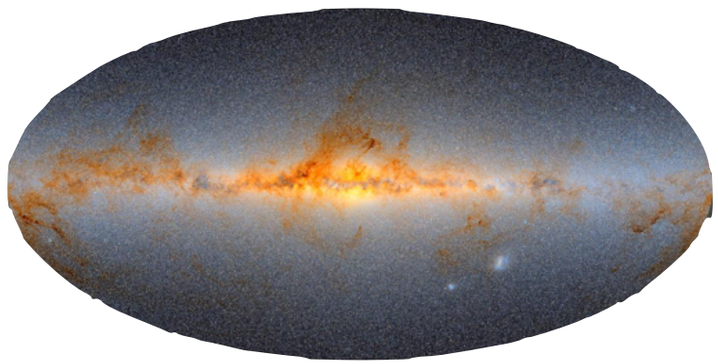
HiPS2FITS
HiPSgen, …

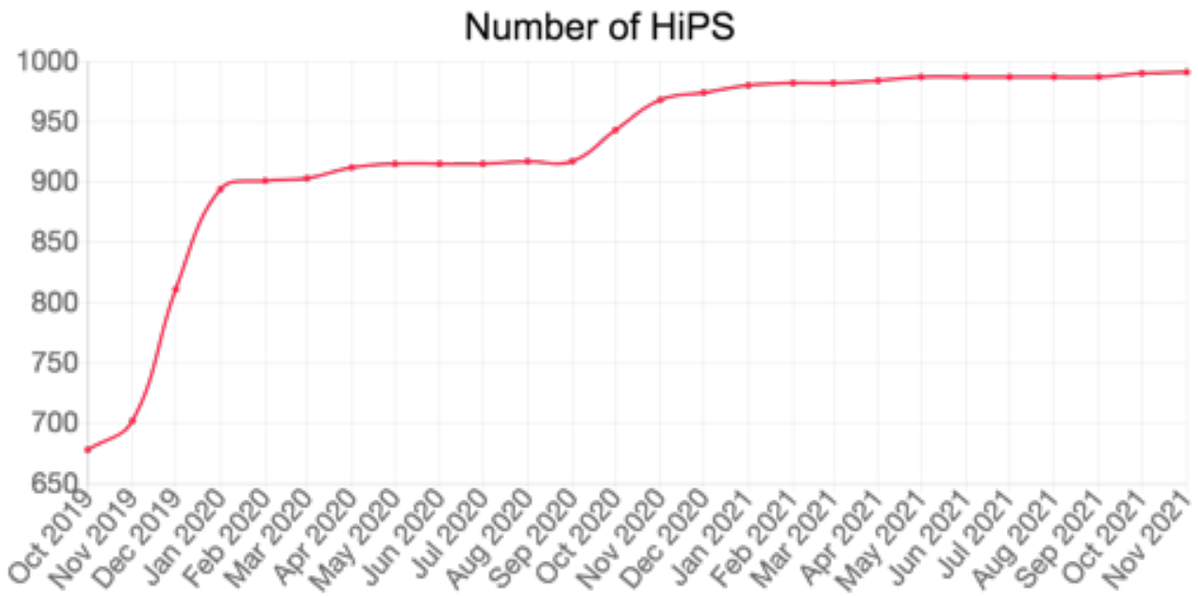


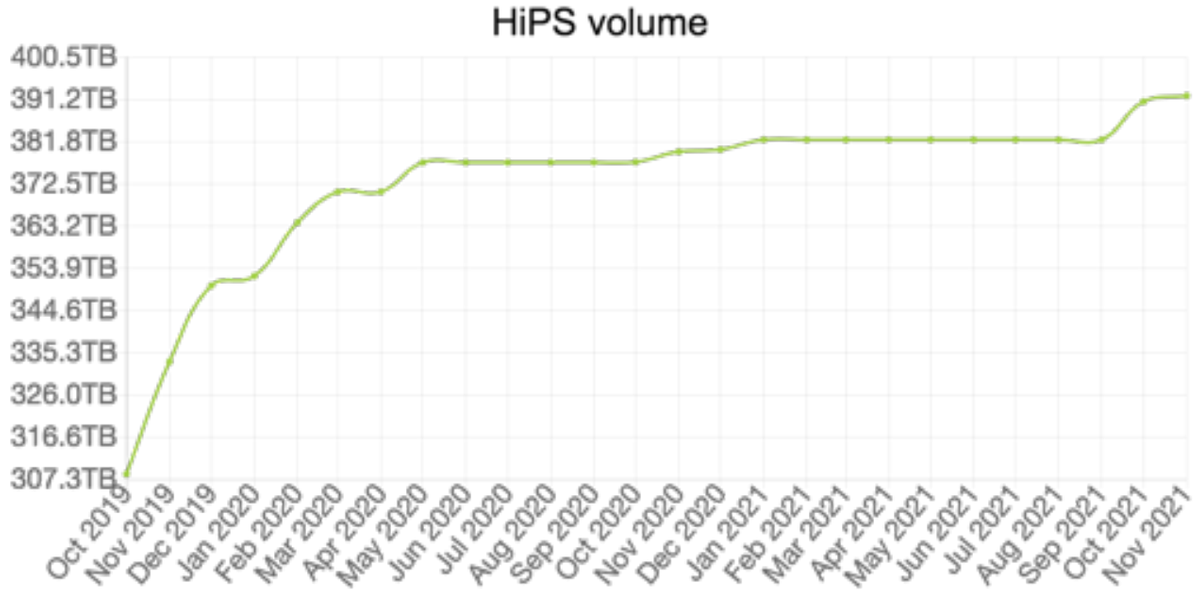
Download
Aladin Desktop
on your machine

Preview with
Aladin Lite
in your browser

HiPS







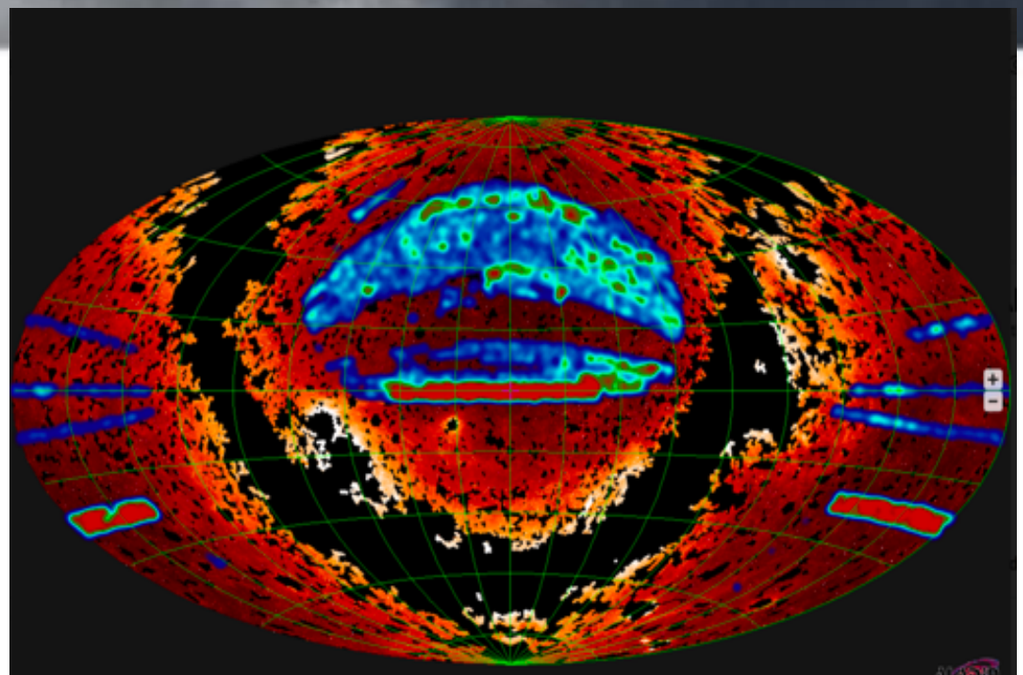
- 47 new HiPS including:
  - MSX, 4 bands
  - Gaia EDR3 flux map
  - ZTF
  - TESS …
- maintenance operations
  - old HiPS have to be done again (e.g. 2MASS)
  - metadata cleaning and update (M. Buga)

Move to Unistra data center

- as planned by disaster recovery plan
- May 19 2021
- move of one of the All-Sky-Data system (part of HiPS storage)

Aladin Lite

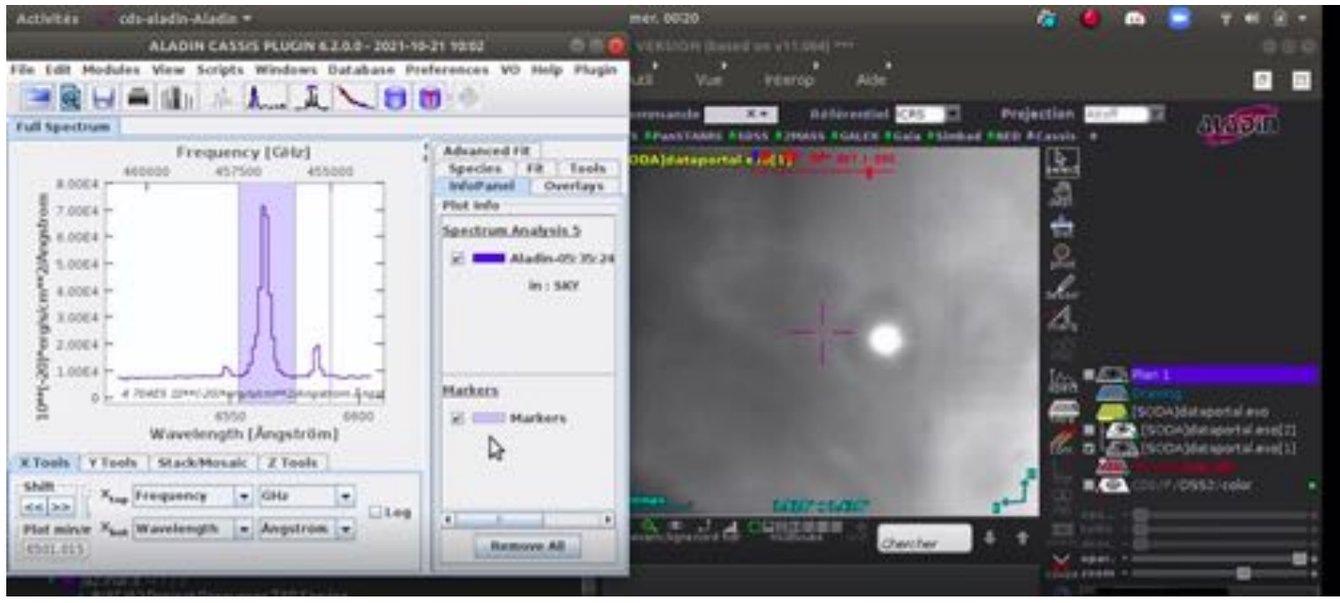
- number of startup from sites embedding Aladin Lite is strongly increasing
    - 25k startup per day (+71%)
- Many new functionalities developed by Mathieu Baumann as part of ESCAPE project. Use of WebGL/WebAssembly
    - support multiple projections
    - support HiPS FITS tiles
    - density map catalog visualizations
    - improved rendering
- Aladin Lite V3 coming soon



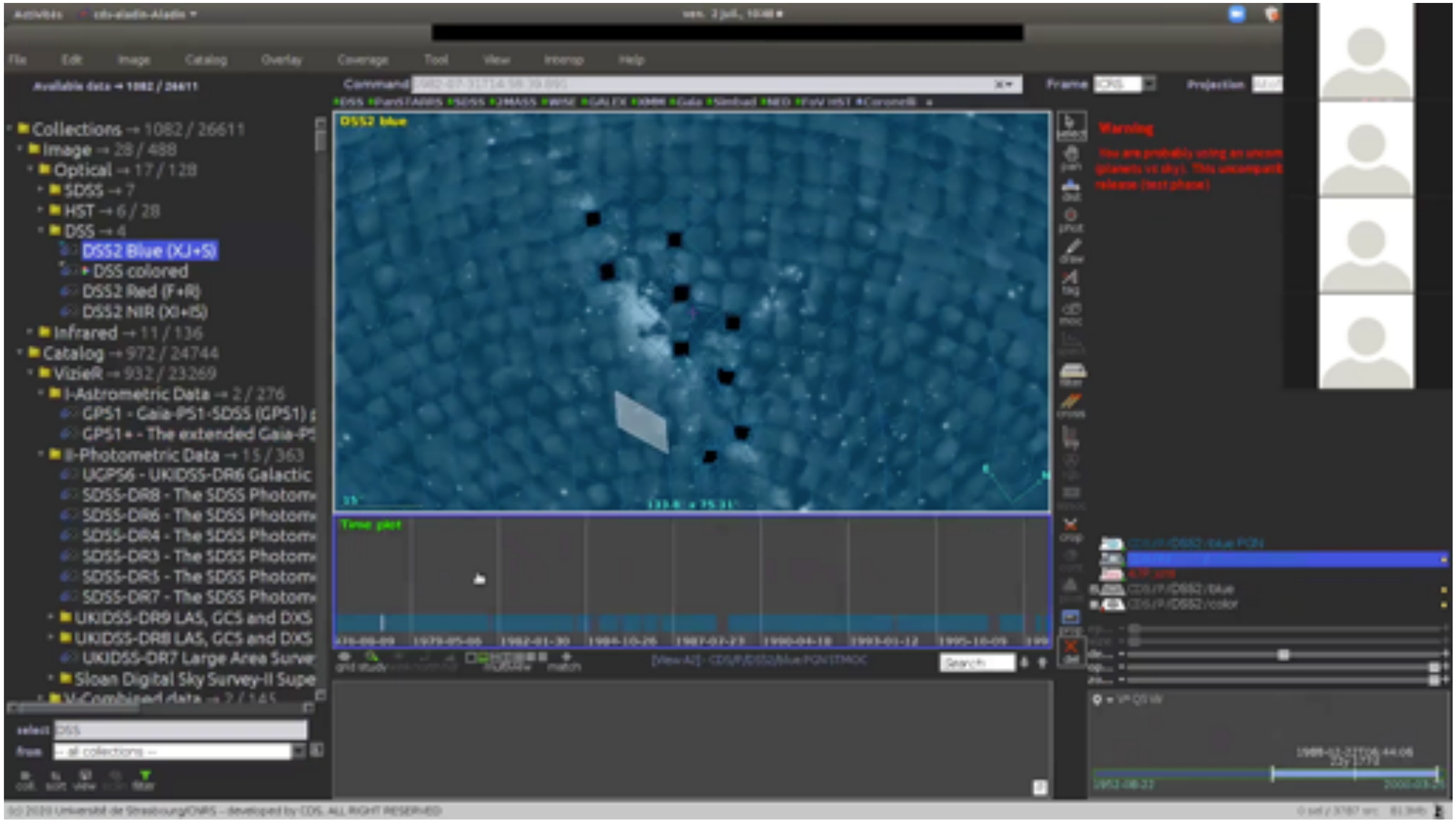
I/313/lqrf sources catalog overlaid on GALEX HiPS tiles

Aladin Desktop

- Since V11:
  - time domain capabilities
  - now prototype enabling search with time constraints
- Statistical pixel analysis capabilities
  - CASSIS plugin for Aladin
  - presented at ADASS, good user community feedback
- Aladin V12 planned for spring 2022

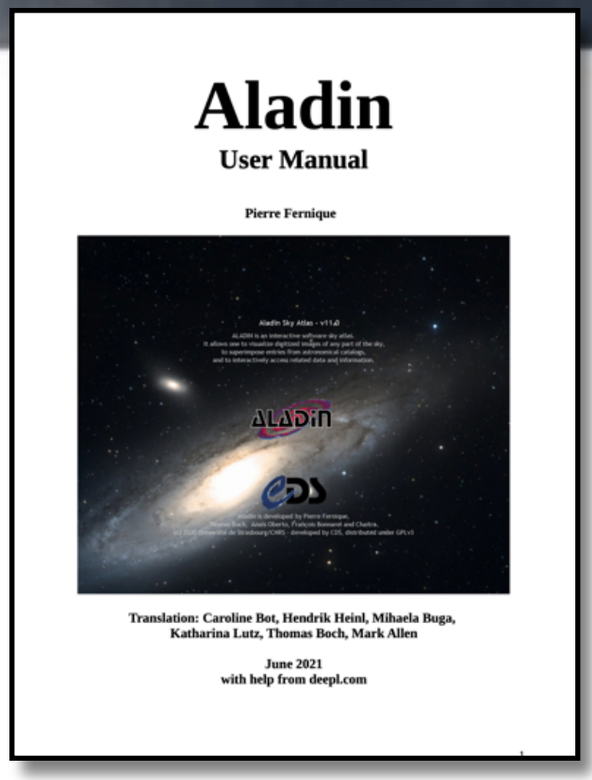


Aladin Desktop / Time domain demo



Interaction with user community

- Aladin manual in english!!
  - coordinated effort from *H. Heinl, C. Bot, M. Buga, K. Lutz, P. Fernique, T. Boch, M. Allen*
  - V10 ready for ESCAPE VO school, soon after for V11


Aladin

User Manual

Pierre Fernique

Translation: Caroline Bot, Hendrik Heinl, Mihaela Buga, Katharina Lutz, Thomas Boch, Mark Allen

June 2021
with help from deepl.com

- Conferences, workshops, school
  - ADASS (4 papers), AAS, China VO conference, ESCAPE VO School
- Case-by-case assistance
  - LIGO/VIRGO project
  - NOIRLab + ESA/Hubble outreach pages

Planetaria and HiPS

- New interest in HiPS for planetarium sky simulators
  - Digistar 7
  - RSA Cosmos
- Will be implemented in new Strasbourg planetarium
- Workshop “Astronomie en partage” on scientific mediation in planetarium thanks to research data

Future

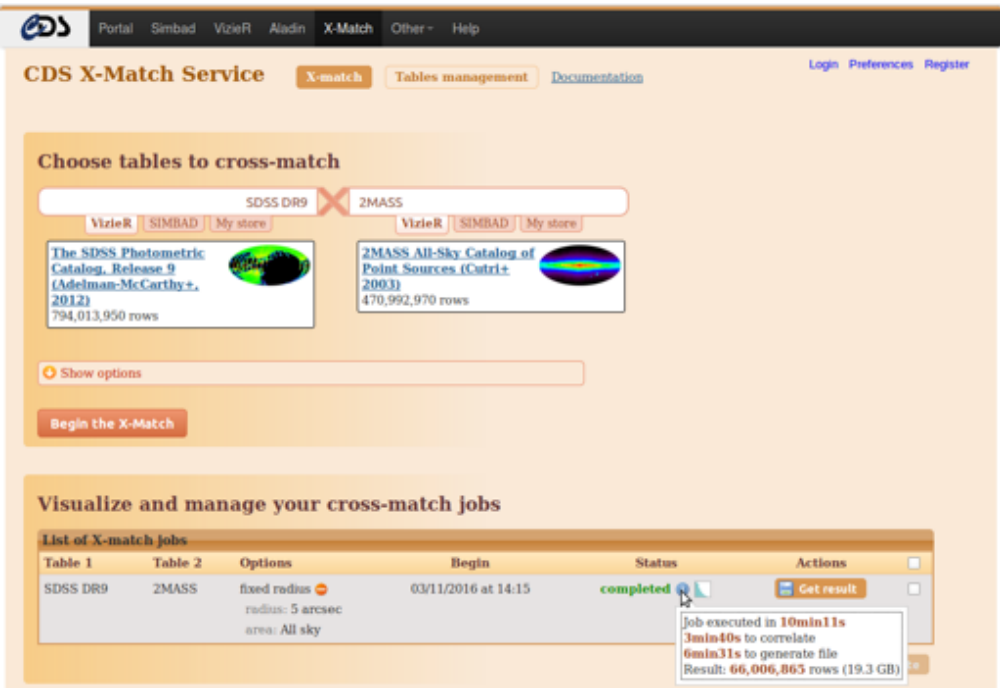
- Aladin Lite V3
- Aladin Desktop V12

coming soon!

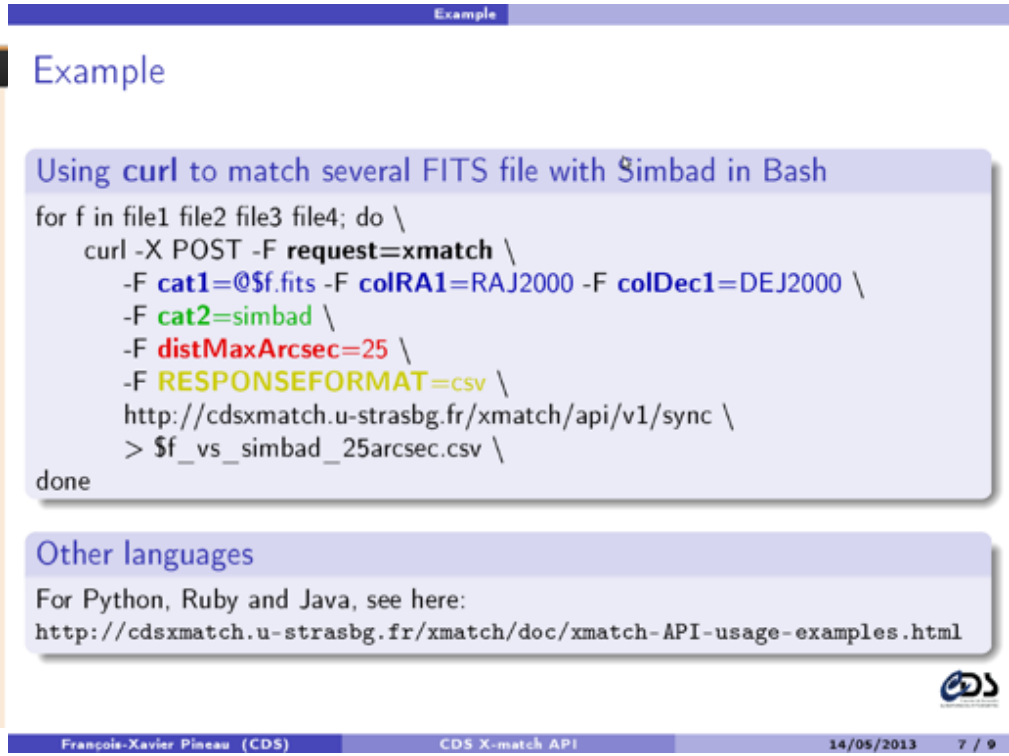
- Strategy
  - Jupyter notebooks
  - Renewal of the All-sky data storage (2025-2026)
  - Development of new services centered on HiPS
  - …

CDS XMatch Service

- Very efficient cross-match of two (possibly large) tables
  - Any VizieR table and Simbad
  - User uploaded table
- Smooth operations in 2020-21 with current software
- We doubled the disks capacity

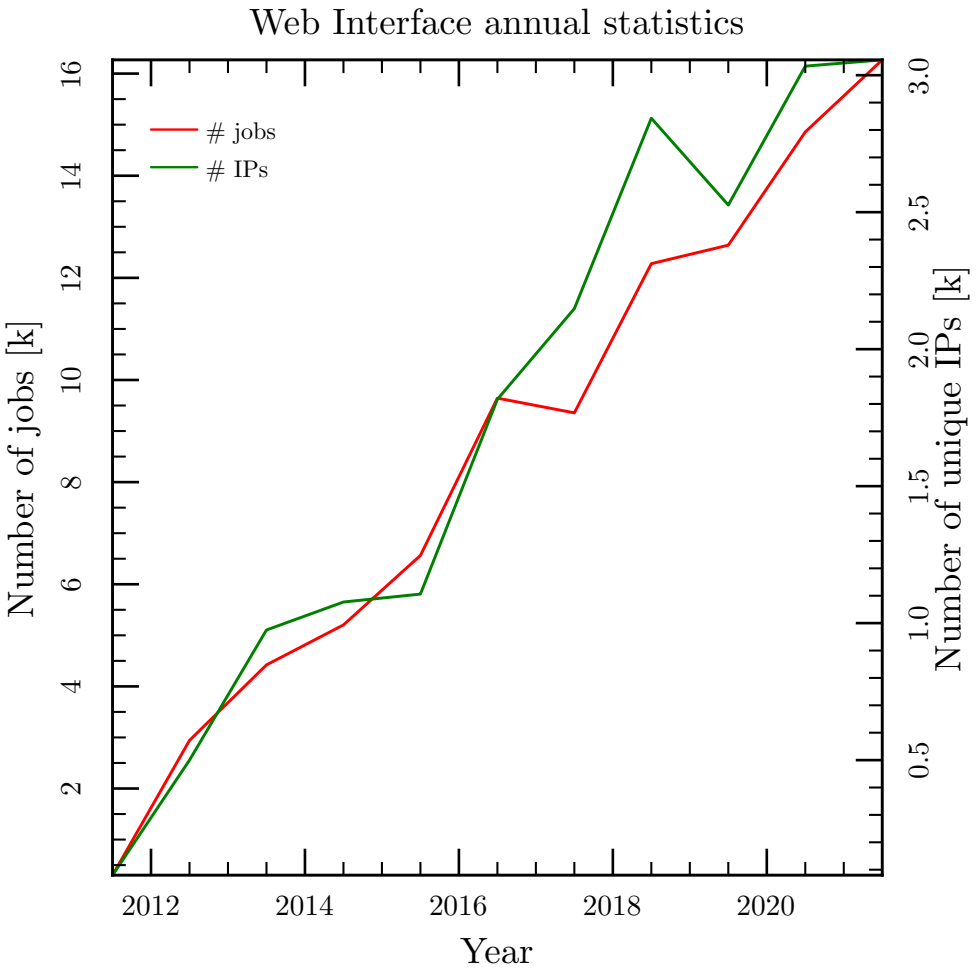


Web interface



Programmatic access (HTTP API)

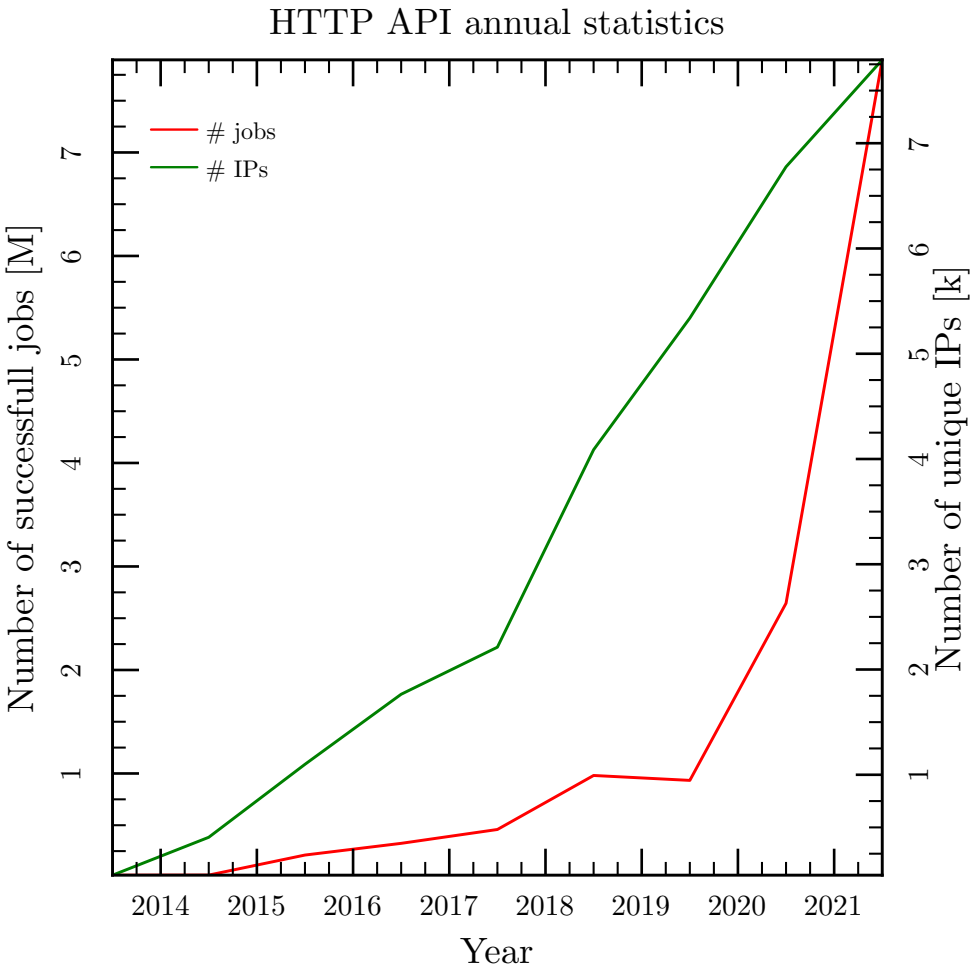
Annual usage statistics



Web interface

- #nJobs still growing
- Produces $>$**70** million associations / day

Annual usage statistics



HTTP API (synchronous jobs)

- Usage still growing too
- #nJobs: $>$**5000** jobs/day (+100%)
- #nIPs: $>$ 900/month
- $\approx$**80** million associations / day
- But pics of activities by a low number of users

On-going developments

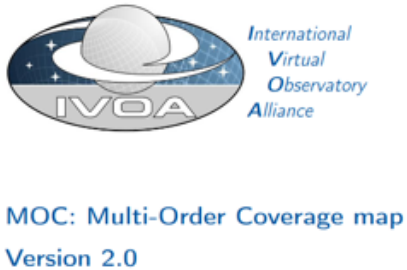
- 2018:
  - Start Rust (+WebAssembly) evaluation
  - Port the CDS kd-tree code from Java to Rust (and WebAssembly)
- 2019:
  - CDS Healpix Library
  - New cross-match engine prototype (ExXmatch): presented at ADASS
  - Start of a Serialization/Deserialization engine
- 2020:
  - Continue Serialization/Deserialization engine
  - Expression evaluation library (SQL SELECT/WHERE like functionalities)
  - Generic catalogue interrogation (VizieR large cats + Xmatch)
    - ⋆ column selection (reduce the number of output columns)
    - ⋆ post-filtering (reduce the number of output rows)

On-going developments

- 2021: Start an astrometrical Rust library
  - done:
    - ★ proper motion propagation
    - ★ parsing / on-the-fly generation of IAU names
  - to be done: reference frames conversions
  - purpose
    - ★ better understand and support the existing C and Java libraries (F. Ochsenbein)
    - ★ allow queries/xmatches taking into account proper motions

On-going developments

- 2021: MOCLibRust a common lib for MOCPy, MOCCli and MOCWasm
  - No directly xmatch related, but will be used in xmatch
  - Framework: both the ESCAPE project and the validation of IVOA standards
  - From a single common core library in a single language (**Rust**) we developped:
    - ⋆ a Python module (to target astronomers)
    - ⋆ a library for Web Browsers (e.g. for Aladin Lite, GW portal)
    - ⋆ a command line tool pre-compiled to target several architectures
    - ⋆ the library itself can be directly used in Rust projects
    - ⋆ (possible C, PSQL, ... wrappers)



- We plan to continue this **strategy** for the new xmatch service developments: write a **single code usable in multiple languages / platforms**, on our **servers**, **web applications** or **users machines**.

On-going projects

- Collaboration with GALECOS/XMM SSC members to build SEDs distributed by ESA: `https://www.cosmos.esa.int/web/xmm-newton/xsa`
- Participation to the European project XMM2Athena
  - WP2: provide cross-matches to other WPs
  - rely on CDS expertise (ARCHES projects, ...)
- FINK and Vera C. Rubin brockers
  - participate to brockers meetings
  - support and prospects for using CDS tools (FINK is already using the xmatch service)
  - Scalability issues to be analysed

R&D @ CDS

and various developments

André Schaaff on behalf of the CDS Team

CDS Scientific Council 2021

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

IVOA

ESCAPE

European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

Foreword

- Technological evolutions are fast & various (interactions, visualization, mobility, components, Big & Open Data, Clouds, etc.) with many actors in both commercial and Open Source domains.

- The R&D activity is well identifed, structured and involves several persons of the staff with the help of interns and short contracts.

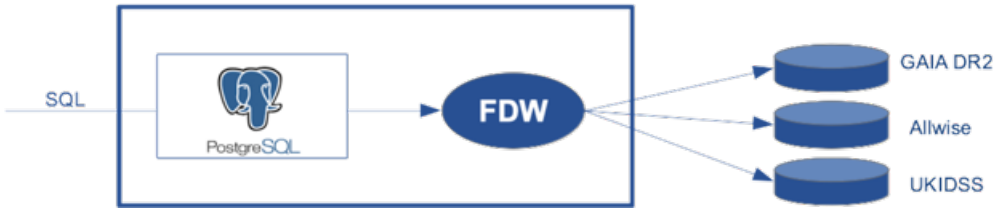
- A continous training through the presentation of the results.

Internship programme

- (Only) 6 interns in 2021 worked with us on a few topics, R&D and various developments
- In 2020 12 (3 years FTE) in telecommuting, 2021 was a transition year and we expect the return to the usual level in 2022
- + short Summer contracts to push the work on the production side.
- A way to hire engineers on projects (in a tight IT Job Market).
- Remark: not exhaustive, other developments and experiments have also be done.

VizieR Access to remote large tables

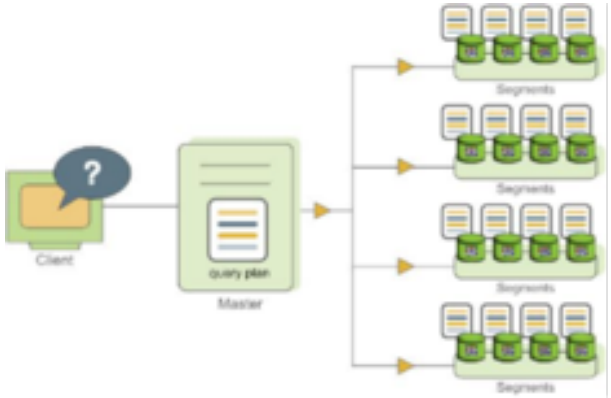
- A mid-term R&D with in 2019/20 an experiment around remote large table access in databases
- Integration of large tables stored in CDS-binary format wrapped in PostgreSQL database using the Foreign Data Wrapper technology.
- a possible solution for very large tables: Euclid, LSST...



G. Landais, G. Mantelet, F.-X. Pineau
Interns: Alexandre Lepicier (UTC) in 2020
Adrien Vizier-Fontanesi (ENSIIE Paris) in 2019

VizieR Access to a distributed database

- Followed in 2021 by an experiment around a distributed database for large tables
- Exploration of the GREENPLUM technologies, based on PostgreSQL, which enables to distribute large tables on several database nodes



GREENPLUM DATABASE

G. Landais, G. Mantelet, O. Marchal
Intern: Yann Miglionico (IUT Schuman Illkirch)

VOSpace

- A prototype was developed in 2020 by the same intern, Grégory Adam
- In 2021, he worked on its performance improvement with Apache JMeter
- Outcomes: ok for the prototype, not yet tested on other VOSpace implementations

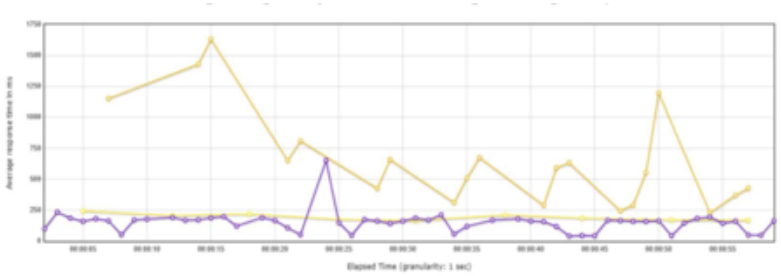


Figure 16 : Temps de réponse moyen en fonction du temps pour des requêtes sur les fichiers avec un utilisateur faisant des requêtes toutes les secondes, un autre toutes les 100ms et 7 utilisateurs faisant des requêtes toutes les secondes

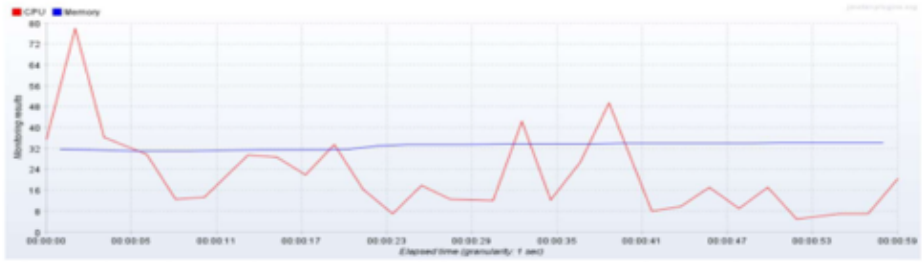


Figure 18 : Utilisation du processeur et de la mémoire en % en fonction du temps pour un scenario mixte

G. Mantelet
Intern: Grégory Adam (TPS Illkirch)

VOSpace, to a micro-service version

- Developped by Grégory Adam in the frame of a short contract

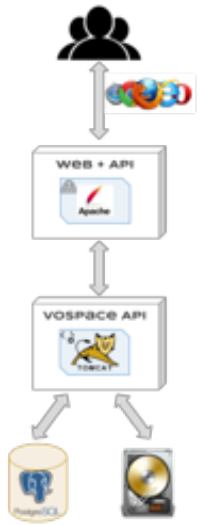
Technologies:

- Docker *(container system)*
- Quarkus *(Java + Netty)*
- NginX *(web server)*
- MongoDB *(NoSQL database)*

Outcome:

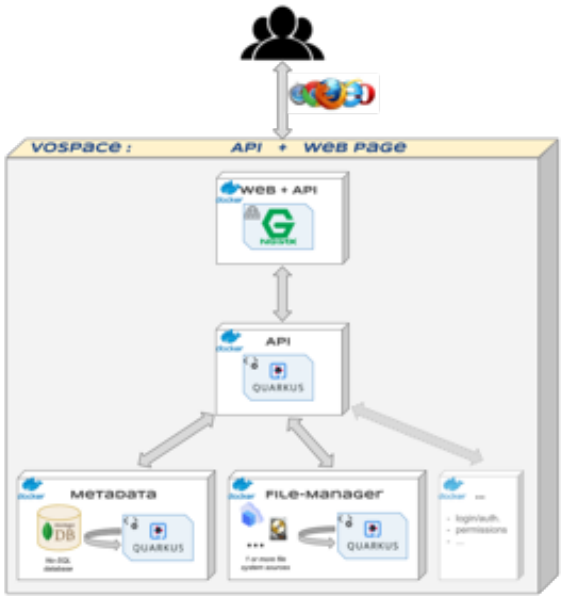
- Good functional prototype
- Possibility to use with other CDS services
- Need additional work especially on the VO and integration aspects

First Prototype



- Monolithic
  - Big application sources
  - Heterogeneous sources
- A failure impact all the application
- Heavy database for key/value data
- Installation difficulties (all tools must exist and configured)

New Prototype



- Modular
  - One application for each task
  - Simpler and shorter sources
- A failure impact only one application
- Faster database
- Able to deal with multiple file systems
- Easier and faster to deploy thanks to Docker

G. Mantelet
Short contract: Grégory Adam

Annotating TAP responses on the fly with IVOA data models

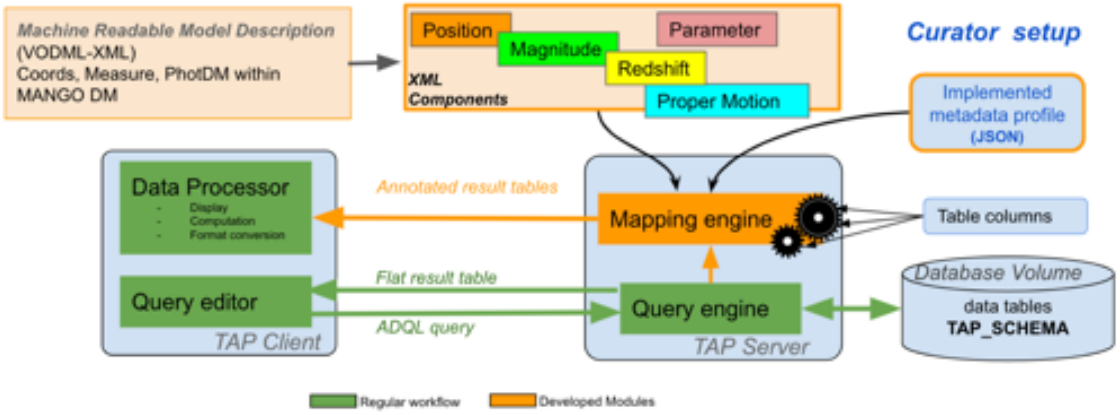
- Goal: Define a way to annotate a VOTABLE TAP response with IVOA data model information
- Test case deals with catalogs of astronomical sources
- **Coords, Measure, Photometry data models** propose classes to bind the various fields from the tables into attributes.
- The **MANGO data model** (work in progress) aggregates these features within an object, with one source identifier and some associated data when necessary.

F. Bonnarel, M. Louys , L. Michel
Intern: Joann Vetter (ENSIIE Evry)

Annotating TAP responses on the fly with IVOA data models (2)

- The description of model elements is represented and embed in XML within the VOTable TAP response to be compliant with TAP output

Annotation scenario



BoF at ADASS 2021

TAP and the Data Models

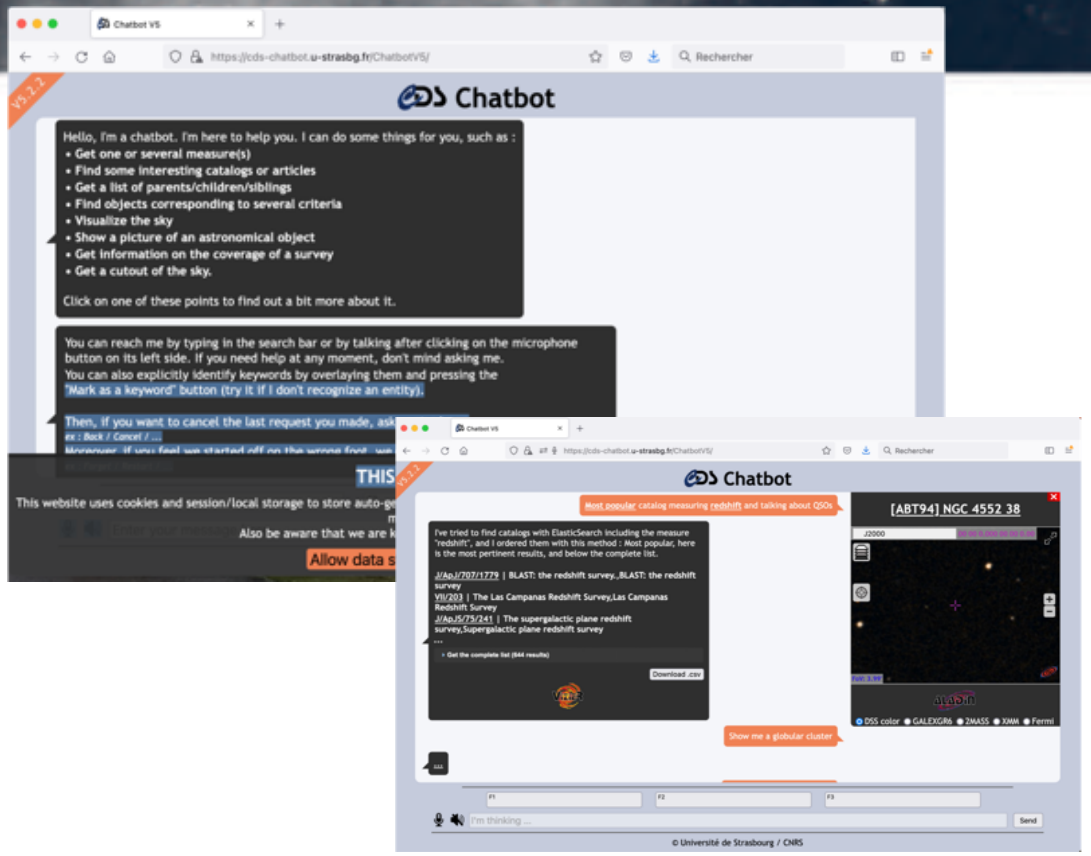
F. Bonnarel, M. Louys , L. Michel
Intern: Joann Vetter **(ENSIIE Evry)**

Chatting with the services

What is the effective temperature of Sirius ?

- A long-term work started in 2017, now again based on an Open source plateform (RASA)
- We focus the effort on the Natural Language translation to understandable queries by the CDS services and the presentation of the results.
- The voice integration opens also a new investigation area, the equal access to the services and the to Research data, especially for people with disabilities.

Show me a globular cluster



It benefits from authors in Simbad, missions and wavelengths in VizieR, DJIN to recognize identifiers in a text, UCDs, ADQL / TAP, the Sesame name resolver, Aladin Lite, => in-house and VO effort

A. Schaaff, T. Boch, S. Derriere. M. Brouty
Intern: Ahmet Adam (UTBM)

Inclusive CDS

- A SARA 2020 topic was Getting everyone connected: software and hardware interfaces that make a difference for professionals and students with disabilities in on-line work; making **online accessible to** everyone around the World.

- IVOA huge work is one of the keys of the Accessibility… but in the sense of FAIR !

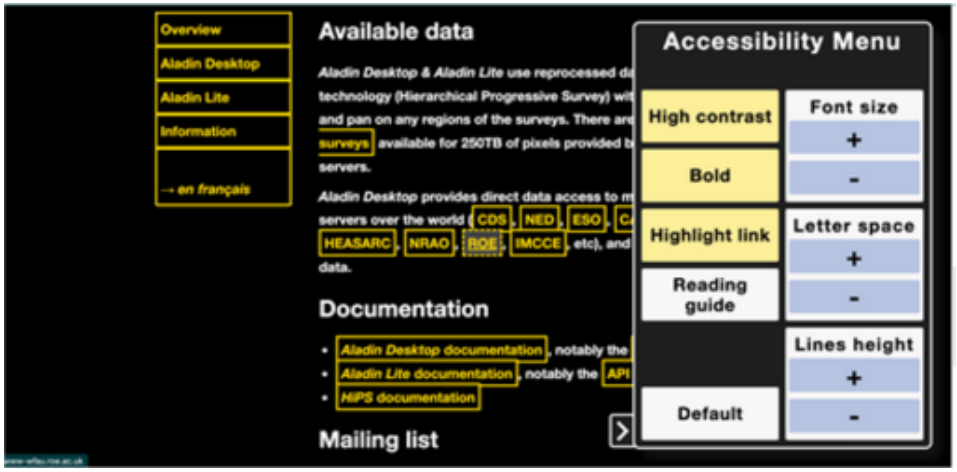
- It sounded great to introduce a noble human concern in addition to the usual Scientific and technological concerns. And to start looking for solutions to make it happen.
- We may already have some of the keys for the Accessibility in the sense of "to all" (We are not starting from scratch).

A. Schaaff, G. Mantelet, T. Boch
Intern: Romain Berger (UTBM)

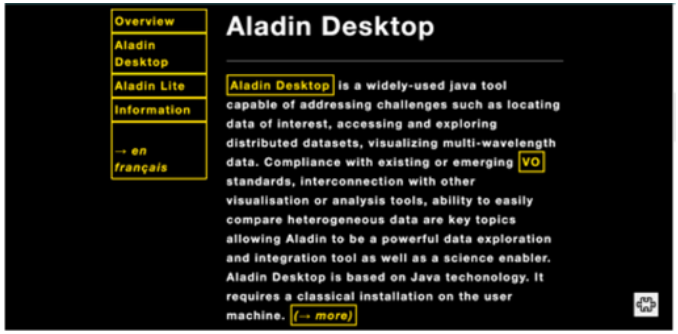
Inclusive CDS (2)

- A six month study started in September 2021 to
  - learn about disabilities
  - draw up an initial inventory of the accessibility of CDS data and services for people with disabilities
  - find solutions to improve the existing
  - set up upstream new ways of presenting data and services, developing interfaces and tools
- Towards an inclusive CDS ?
- Unrealistic and too hard ? Hard, yes but it is a first step
- First experiments are ongoing with CDS webpages, AladinLite, …

Exemple for view accuracy disability



Other experiments with readers to help blind people with a voice translation of the text



A. Schaaff, G. Mantelet, T. Boch
Intern: Romain Berger **(UTBM)**

Future investigations

- As usual not Geek-driven, a continuous R&D effort to provide an added value to the data access & presentation mechanisms.
- Continous work on Science platforms
- A Lab part on CDS Website
- Efforts to make the data and services accessible to Everyone.

Conclusion

- A (less than usual) coverage with various spin-offs:
  - improving the services <- R&D -> updating the staff skills
  - new (but is not enough by itself !) technologies.

- A team work (@ ObAS level): contracts, dedicated pool of workstations, presentation of the services and the professions to integrate quickly the newcomers, etc.
- Back to a larger R&D panel in 2022 !

Science @ CDS

Scientific diversity at CDS

- **Information discovery and processing**

  Ontology, semantics, resource discovery — S. Derrière, F. Genova

- **Stellar astronomy**

  Circumstellar matter, AGB/post-AGB stars, surveys — C. Loup

- **Galactic astronomy**

  Binary stars, X-ray surveys — A. Nebot

  ISM, star formation — L. Cambrésy

  Galactic disk formation, surveys, simulations — A. Siebert, G. Monari

- **Extragalactic astronomy**

  Epoque of reionization, nearfield cosmology — P. Ocvirk

  AGNs, VO science — M. Allen

  ISM, nearby galaxies — C. Bot, Y. Stein, K. Lutz

  Cluster galaxy evolution, gal. centers, GPS sources — B. Vollmer, S. Amodeo

Multi-wavelength astronomy

• X-ray	A. Nebot
• Visible/NIR	C. Loup, S. Derrière, G. Monari, M. Allen
• Optical spectra	A. Siebert, P. Ocvirk, S. Amodeo
• Infrared	C. Bot, L. Cambrésy
• Radio/mm	B. Vollmer, Y. Stein, K. Lutz, S. Amodeo

Scientific responsibilities

• SIMBAD	**C. Loup**, L. Cambrésy, A. Siebert, C. Bot, A. Nebot, P. Ocvirk, B. Vollmer
• VizieR	**P. Ocvirk**, C. Bot, G. Monari
• Aladin	**C. Bot,** A. Nebot
• Nomenclature	**B. Vollmer**
• Dissemination	A. Nebot, C. Bot, S. Derrière, S. Amodeo, K. Lutz
• Special operations	A. Siebert, L. Cambrésy, Y. Stein (SPECFIND)

The role of CDS scientists

- part of integrated team together with documentalists and engineers
- selection and validation of scientific data
- validation of the data description
- scientific advice for / supervision of the development and evolution of CDS services

Scientific input

All CDS scientists contribute their

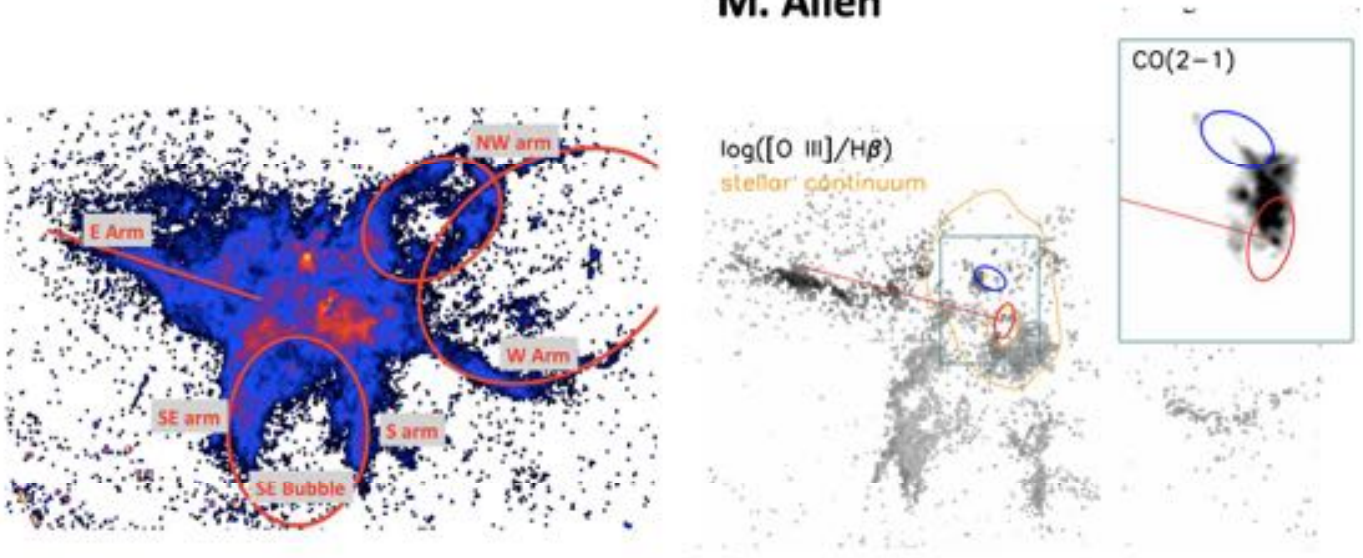
- expertise
- view on astronomy and information science
- experience with CDS services
- new ideas

All CDS scientists contribute to

- the daily workflow/data ingestion
- the development and evolution of the CDS services

Recent result for ionization of gas in active and merging galaxies

M. Allen



Tracing the Ionization Structure of the Shocked Filaments of NGC 6240
Medling et al. 2021 – accepted ApJ (arXiv:2111.01025)

Finalisation of long term project from HST WFC3 observations obtained in cycle 19 (Led by Kewely & Allen). Ionization structure investigated with high-res emission line maps (HST, Keck, ALMA). The 'bubbles' in the gas-rich merger galaxy NGC 6240 show a clear shock front traced by high [O III]/Hβ and [O III]/[O I].

Stefania Amodeo

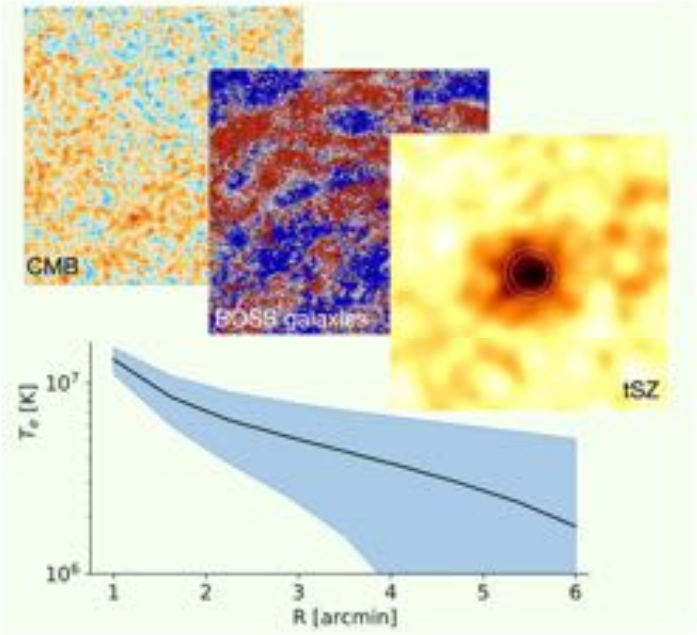
ESCAPE

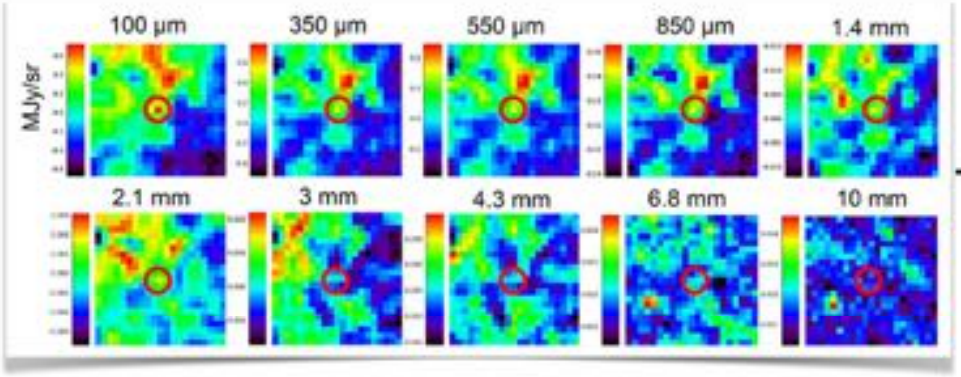
Scientific support to the ESCAPE project (WP4-CEVO)

Notebook tutorials to facilitate *FAIR* science

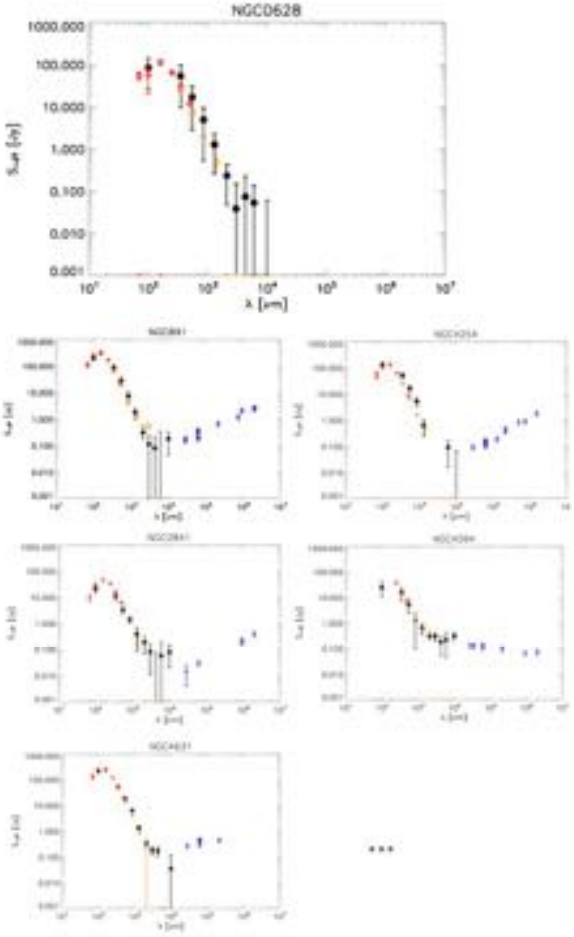
Gas thermodynamics and feedback in galaxy clusters from Sunyaev Zel'dovich effects

Amodeo et al. 2021, Phys. Rev. D, 103, 063514





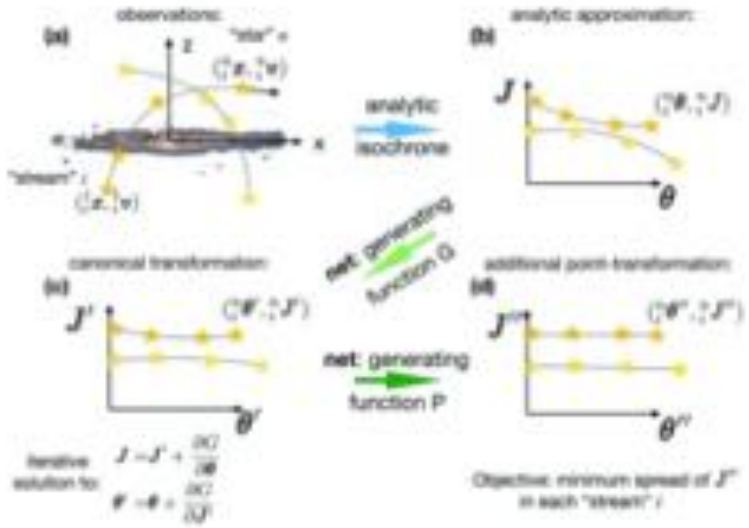
During his internship with **C. Bot & K. Lutz**, using IRAS and Planck, Axel Rymar removed significant foreground and background emission to build a set of spectral energy distributions of 21 galaxies including the microwave regime



G. Monari

The ACTIONFINDER:
An Unsupervised Deep Learning Algorithm for Calculating Actions and the Acceleration Field from a Set of Orbit Segments

Ibata, Diakogiannis, Famaey, Monari (ApJ, 915, 5, 2021)



Through a two steps process, from only the "segment" of one or more orbits (for example positions and velocities of stars in thin streams, observed in the stellar halo of the Milky Way), we are able to determine to a high degree of accuracy both the orbital actions $\boldsymbol{J}$ (integrals of motion, conserved along the orbits) and the gravitational field where they "live" (i.e., the gravitational field of the Milky Way).

Observing the sky in X-rays allows us to detect the hottest and most energetic phenomena in the Universe.

WP2: Multi- wavelength/ multi-messenger counterparts lead by **CNRS/Observatoire Astronomique de Strasbourg**
Objectives: Provide multiwavelength statistical identifications and spectral energy distributions of XMM-Newton X-ray sources including those discovered in the stacked catalogue and new transient sources found by the variability work package. Train to identify electromagnetic counterparts of multi-messenger events.

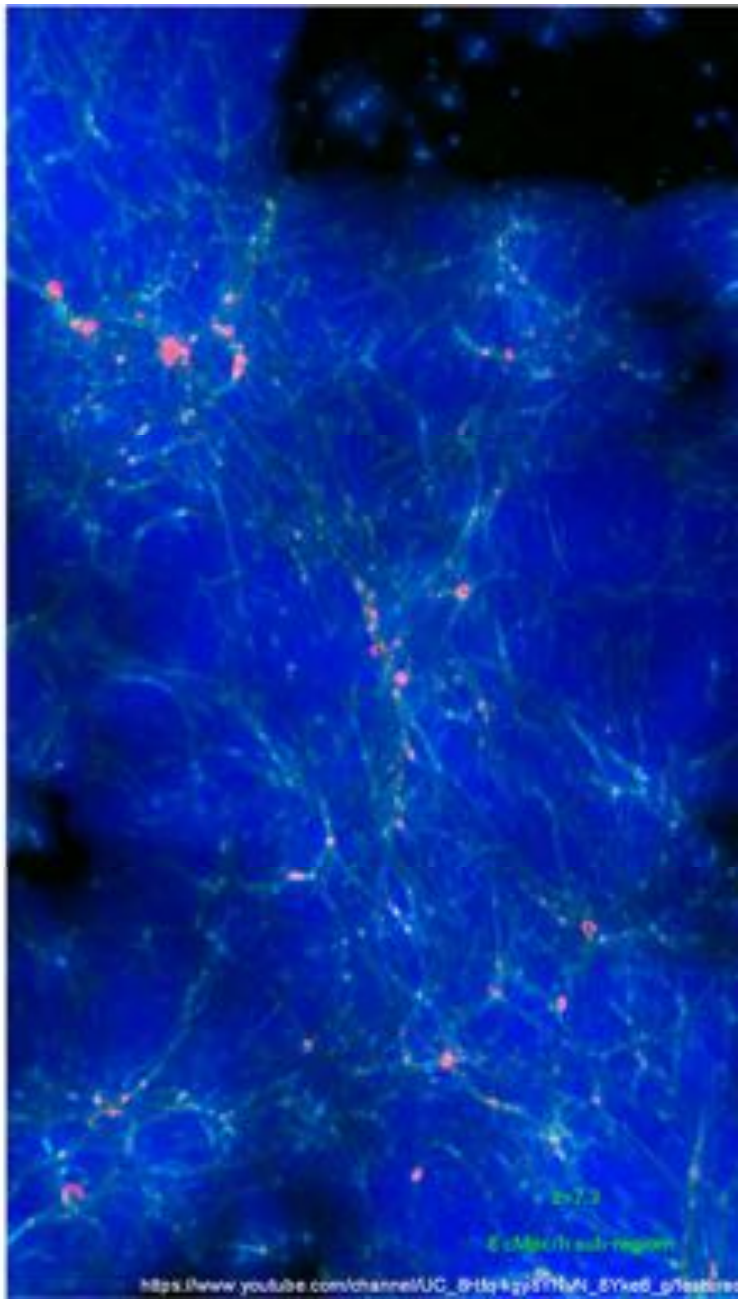
CDS : **Ada Nebot (local coordinator), Francois-Xavier Pineau**, Jere Kuuttila (postdoc) ObAS: Christian Motch, Laurent Michel, Pierre Maggi

http://xmm-ssc.irap.omp.eu/xmm2athena/

P. Ocvirk

Galaxy Formation during the Epoch of Reionization

- PI of Cosmic Dawn III, Radiation-hydrodynamical sim of galaxy formation during the EoR, completed in Feb. 2021 on Summit (Oak Ridge National Laboratory). Topics include:
  - Galaxy populations, impact of rad. feedback, high-z luminosity functions
  - Calibration of sub-grid models for 21cm predictions
  - Lyman-alpha transmission in high-z IGM through EoR
- Produced 20 PB of data in 10 days
- uses Hybrid CPU/GPU computing.
- Member of SKA and NenuFAR science team

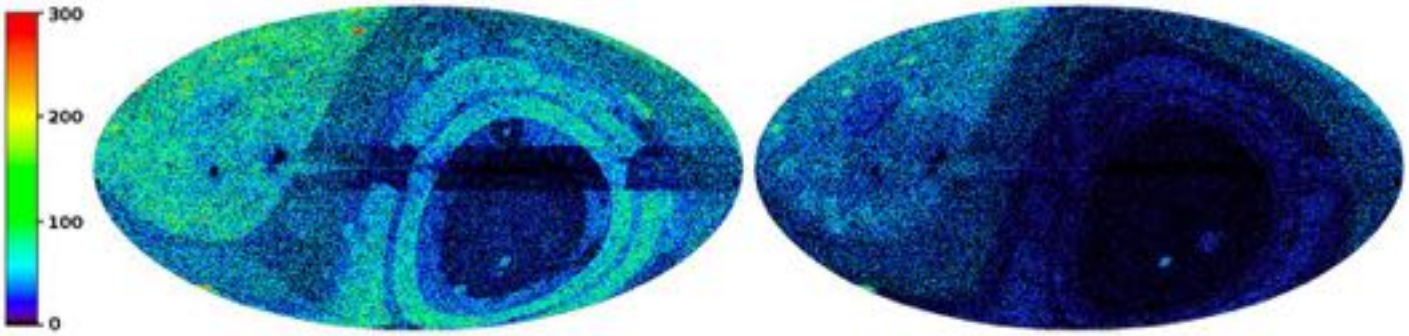


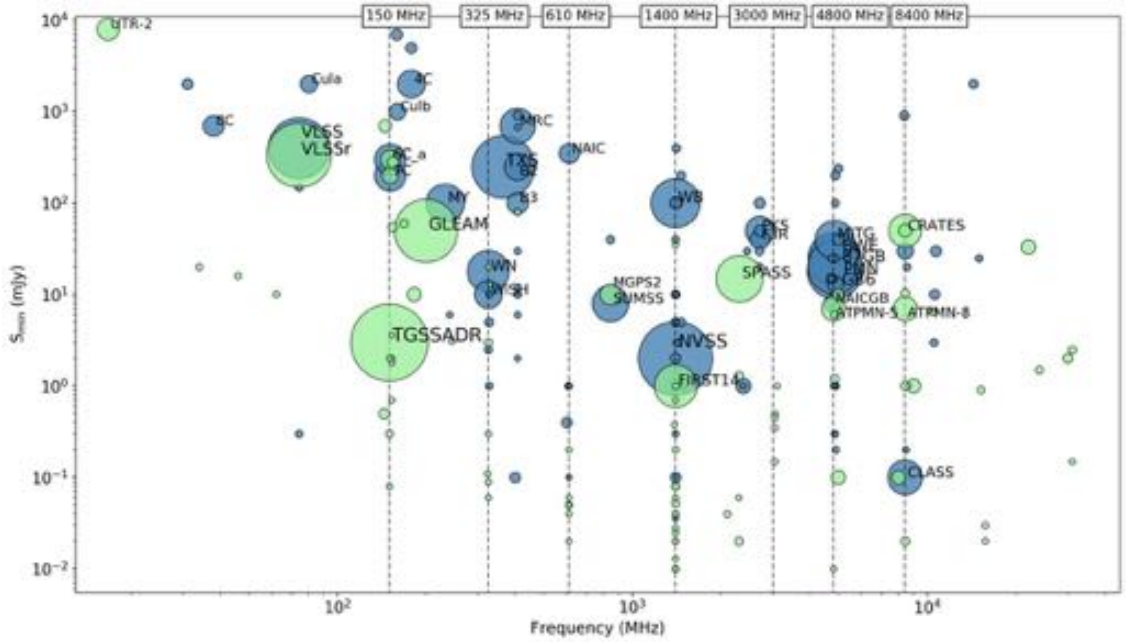
A&A 655, A17 (2021)
https://doi.org/10.1051/0004-6361/202039659
© Y. Stein et al. 2021

Astronomy & Astrophysics

The SPECFIND V3.0 catalog of radio continuum cross-identifications and spectra: Reaching lower frequencies★

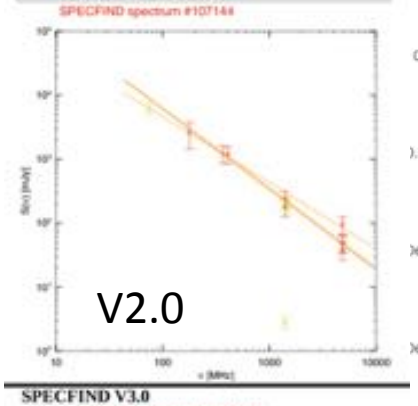
Y. Stein, B. Vollmer, T. Boch, G. Landais, P. Vannier, M. Brouty, M. G. Allen, S. Derriere, and P. Ocvirk

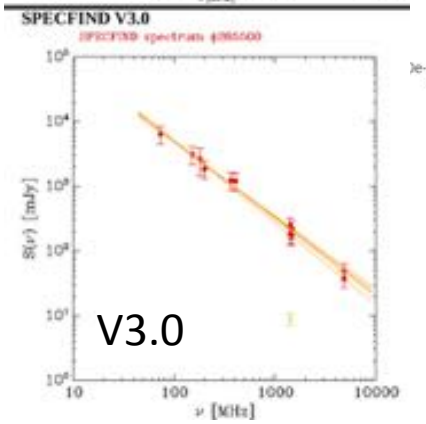


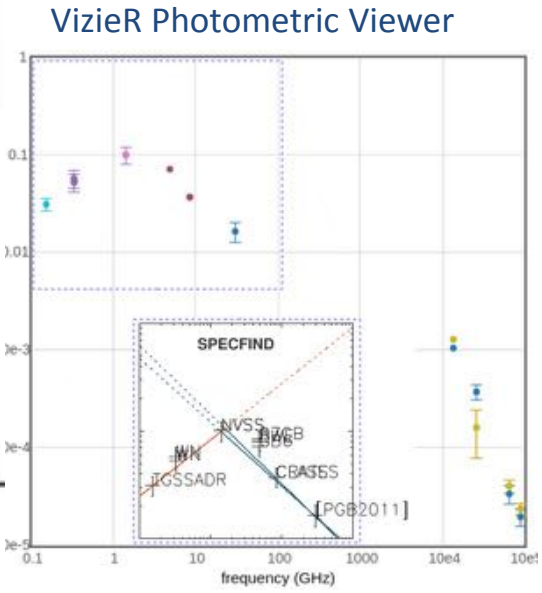


SPECFIND V3.0

- Input tables: 115 -> 204
- Radio SEDs: 107k -> 340k
- Cross-identified sources: 600k -> 1.6M
- Identification of spectral breaks
- Catalog available via VizieR
- Ingestion into SIMBAD planned







Stein et al. (2021)

Conclusion

- Scientists are part of an integrated team
- Scientific diversity is vital for the CDS
- active science @ CDS
- CDS scientists mostly rely on external collaborations
- Independent scientific activities have to be preserved and acknowledged

CDS Plans and challenges

November 30, 2021

Mark Allen - Director CDS

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

CDS Strategy

- **Constantly evolving** — based on scientific needs
- **Main strategy** for core services is well established:
  - Pursue the CDS mission at the highest possible level
  - Science-driven Data Centre for “reference” data
  - Spirit of Open Science, and application of FAIR principles

- **Themes**

i) Reinforcement of core mission — trusted reference data centre
ii) Enabling science with the CDS services — supporting specific scientific projects, direct support of astronomers, and development of the CDS science team
iii) Engagement with the astronomy community
iv) Adaptation and innovation — responding to science needs and increasing volume
v) Building on success of CDS by maintaining specialised staff profiles & teamwork

Plans 2021-2022

Continuation of core work of building CDS content and operating services.

Immediate topics:

- Managing with the current continuing uncertainty.
- Re-grouping to **address the consequences of 2020-2021**.
- Work toward a *"new normal"* with different rhythms.
  - Tele-working (up to 2 days for various roles).
  - Include Virtual conferences/meetings as part of regular operations.
  - Put high value on in-person interactions.
  - A changed approach to travel / missions

Plans 2021-2022

- **CDS Service Developments.** *(From yesterday's presentations)*
  - SIMBAD object types
  - Deployment of BCS (CDS Biblio)
    - Manage with changes in publisher systems
  - Progress on New Sesame
  - VizieR 2 - planning
    - Establish an approach for the very large data sets - answering 'should CDS develop such capacity?'
    - Survey of up-coming missions and surveys
    - Engage a contract engineer for VizieR
  - Progress on the new X-Match prototype
  - Aladin Lite v3.

- CDS Web Pages
- CDS Hotline
- Improvement to CDS code tracking - Gitlab, Github
- Transition to domain name cds.unsitra.fr
- SIMBAD Web interface / APIs
- Jupyter notebooks
- Follow-up opportunities with planetariums

Plans 2020-2021

- **CDS infrastructure**
  - Delays for more moves into the UNISTRA data centre
  - More investment in our server room for electrical redundancy for cooling system
  - Continue investments in hardware - servers / cold-storage
    - Virtual servers and controllers
- Plan for major upgrades: All-Sky-Data ~ 5PB in ~2025
- Continue program of workstation improvements
  - (laptops / desktops / monitors)

Important data sets in 2022

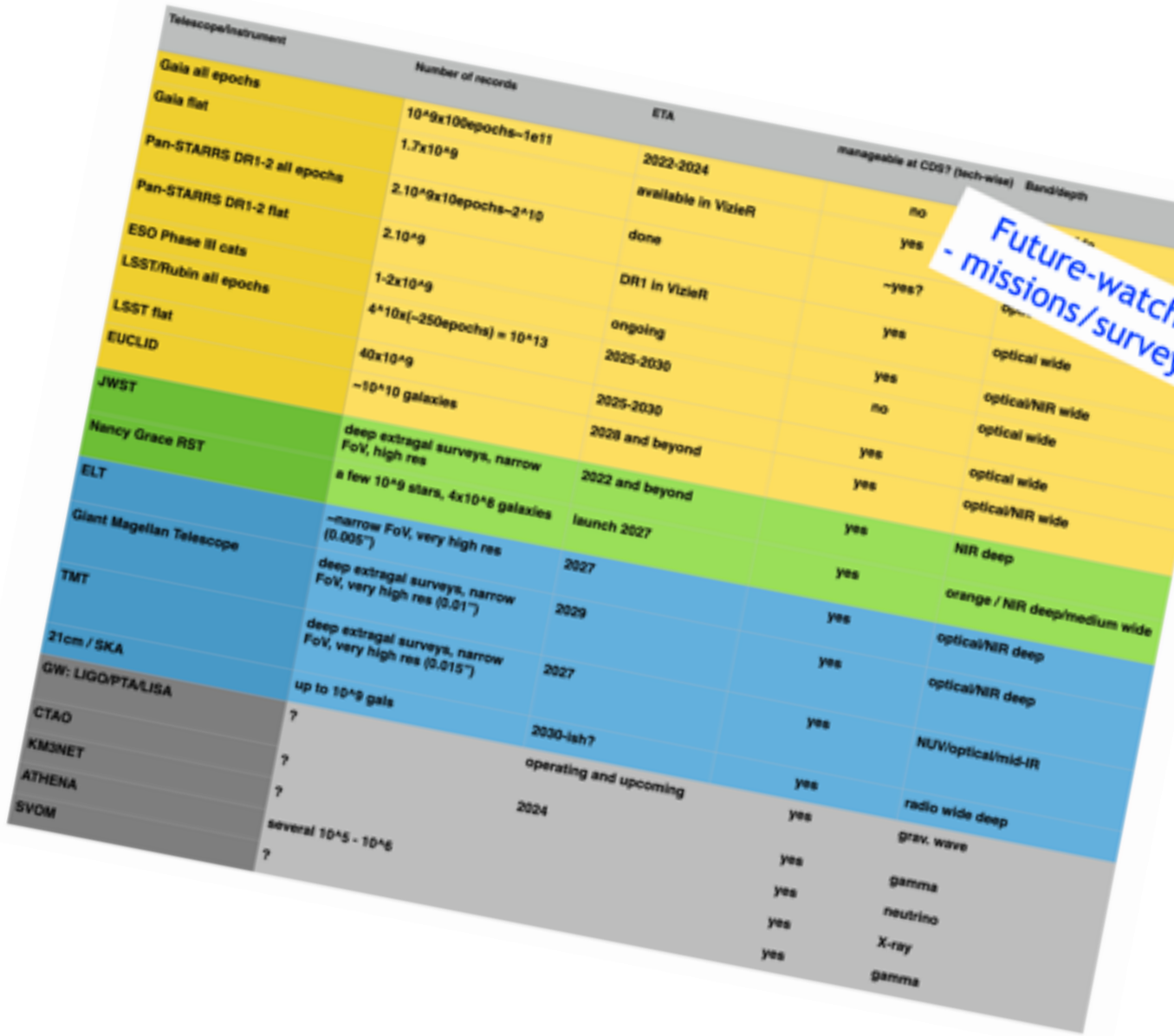
Maintain approach of being as complete as possible with main astronomy journals

What is the “reference data” from future large missions/surveys that should be made available at CDS ?

- ‘Complete’ the future-watch table

2022:

- Gaia DR3
- GSC 2.4
- PanSTARRS DR2
- ESO catalogues
- +++



Telescope/instrument	Number of records	ETA	manageable at CDS? (tech-wise)	Band/depth
Gaia all epochs	10^9x100epochs~1e11	2022-2024	no	
Gaia flat	1.7x10^9	available in VizieR	yes	
Pan-STARRS DR1-2 all epochs	2.10^9x10epochs~2^10	done	~yes?	
Pan-STARRS DR1-2 flat	2.10^9	DR1 in VizieR	yes	optical wide
ESO Phase III cats	1-2x10^9	ongoing	yes	optical/NIR wide
LSST/Rubin all epochs	4^10x(~250epochs) = 10^13	2025-2030	no	optical wide
LSST flat	40x10^9	2025-2030	yes	optical/NIR wide
EUCLID	~10^10 galaxies	2028 and beyond	yes	NIR deep
JWST	deep extragal surveys, narrow FoV, high res	2022 and beyond	yes	orange / NIR deep/medium wide
Nancy Grace RST	a few 10^9 stars, 4x10^8 galaxies	launch 2027	yes	optical/NIR deep
ELT	~narrow FoV, very high res (0.005")	2027	yes	optical/NIR deep
Giant Magellan Telescope	deep extragal surveys, narrow FoV, very high res (0.01")	2029	yes	NUV/optical/mid-IR
TMT	deep extragal surveys, narrow FoV, very high res (0.015")	2027	yes	radio wide deep
21cm / SKA	up to 10^9 gals	2030-ish?	yes	grav. wave
GW: LIGO/PTA/LISA	?	operating and upcoming	yes	gamma
CTAO	?	2024	yes	neutrino
KM3NET	?		yes	X-ray
ATHENA	several 10^5 - 10^6		yes	gamma
SVOM	?			

50
CDS
1972-2022

CDS-50 Schedule:

January - A kick-off webinar announcing the year of events
- Webinar series

February 21-24 ESCAPE VO School to be held in Strasbourg (Hybrid mode)

May - Gaia DR3 - CDS contribution to the ObAS Gaia DR3 events

June 12-16 - AAS Summer Meeting - CDS Booth - 50th anniversary

June - SF2A - investigating whether we can highlight CDS-50 at a session/other

June 27 - July 01 - EAS Annual meeting (Valencia, Spain) - CDS Lunch Session (accepted!)
- CDS booth - 50th anniversary

Aug 16-27 : IAU (Busan, Korea) - CDS booth - 50th anniversary

September 11-16 : ADASS (Victoria, Canada) - CDS booth - 50th anniversay

October / November:
Conference/workshop
- Open science in Astronomy 'from pixels to publications'

(18-20 October?) (followed by school holidays)(Parliament session 17-20)
(8-10 November (followed by 11 November holiday)
(15-17 November)
(22-24 November) (followed by Strasbourg Christmas Markets)(Parliament session 21-24)

Other activities:
- Open a survey on the use of CDS services
- Make use of network to present CDS-50 in seminars at targeted institutes/countries
- Communication materials (short film?)

Open Science in Astronomy – *from pixels to publications*

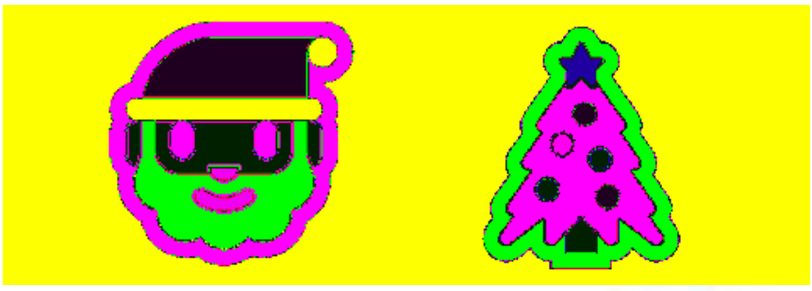
On the occasion of the 50th anniversary year of the Centre de Données astronomiques de Strasbourg (CDS) in 2022, we propose to organise a conference on Open Science in Astronomy. The CDS has been at the forefront of data sharing in astronomy and we wish to celebrate 50 years of progress of the CDS services in a conference that attracts major actors in the international astronomy community, as well as CDS partners and representatives of the French Open science community to look forward to the future plans and challenges of Open Science in astronomy.

The themes of the conference will be developed along the lines of :

- The use of Open Science tools and services for astronomy research
- Current and future reference data sets in astronomy
- Plans and challenges for bibliographic services in astronomy
- Innovations to support scientific discovery with big a rich astronomy data sets
- The FAIR principles in astronomy and beyond.
    - FAIR – Findable, Accessible, Interoperable and Re-usable

Challenges - staffing

- **Administration**
  - CDS cannot work without administration **- long term replacement essential**
    - Contracts, Projects, Reporting, and an increased administrative load
- **Permanent Researchers**
  - 10 researchers — FTE ~3.75 (+1 Emeritus)
    - To be addressed by CNAP and CNRS candidates, & mobility (?)
- **Documentalists -** increased load needs to be addressed
  - Contract positions being opened in 2021/2022
  - **Replacement for the departure of a documentalist is essential**
- **Software Engineers**
  - **Contract engineer recruitment for Vizier URGENT**
  - **Research Engineer retirement - 2024 !! replacement needed on this timescale**
  - Research engineer (bibliography) retirement ~2029

Summary

- **A busy year ahead in 2022**
  - Re-grouping to work toward the CDS mission
  - Preparation of the HCERES evaluation (of ObAS)
- Many very exciting things to do in the community
  - A lot of innovations, improved services, and the message of Open Science
    - CDS-50 as a theme throughout the year with specific events
      - Strengthen links with partners
- Projects - working well, helping to prepare the future, and connection with the wider world of data sharing and large astronomy projects
- Staff & team support necessary — work together to meet the challenges