CDS Scientific Council Meeting: Nov 30 - Dec 01, 2020
Virtual Meeting

Agenda

14h Strasbourg, 13h London, 8h Wash. DC, 21h Beijing

Monday 30 November
14h - Welcome and introductions
14h10 - CDS Activities 2019-2020. *(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 - X-Match. (F-X. Pineau)
15h40 - R&D. (A. Schaaff)
15h55 - Integration demo. (S. Derriere)
16h05 - Break
16h25 - Discussion *(All)*
17h - Close

Tuesday 01 December
14h - CDS plans and challenges. *(M. Allen)*
14h30 - **Closed session**
16h - Close

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

Bruno Bezard (INSU representative delegated by G. Perrin)
Pierre-Alain Duc (Dir. Obs. Strasbourg)
Yannick Hoarau (Université de Strasbourg)

CDS Activity Report 2019-2020

November 30, 2020

Mark Allen - Director CDS

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

2019-2020

- **Core work of building CDS content and operating services**

- Special things in this period:
  - Covid-19 since March 2020…
  - Rapid move to remote working
  - Emphasis on operational stability and continuity
  - Adaption of procedures
  - Communications - videocons, rocket-chat
  - Virtual meetings / conferences
  - Disaster Recovery Plan - Service *“survival sheets”*

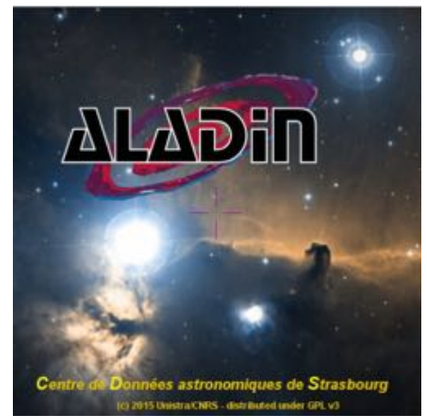
2019-2020

- Releases of innovative new functions
- Improvement/renewal of services/processes
- CDS services continue to be heavily used
  - 1.9 million queries/day *- see next presentation*
- Contributions: IVOA, RDA, EOSC
- Projects — RDA, Europlanet, AENEAS, **ESCAPE**
- Emphasis on interaction - **AAS, ADASS, EAS**, UNISTRA, French community , Time Domain Community, Radio Ast. Community
- CDS Science team progress
- Involvement in networks at national level
- Difficulties - …like everybody, dealing with the pandemic

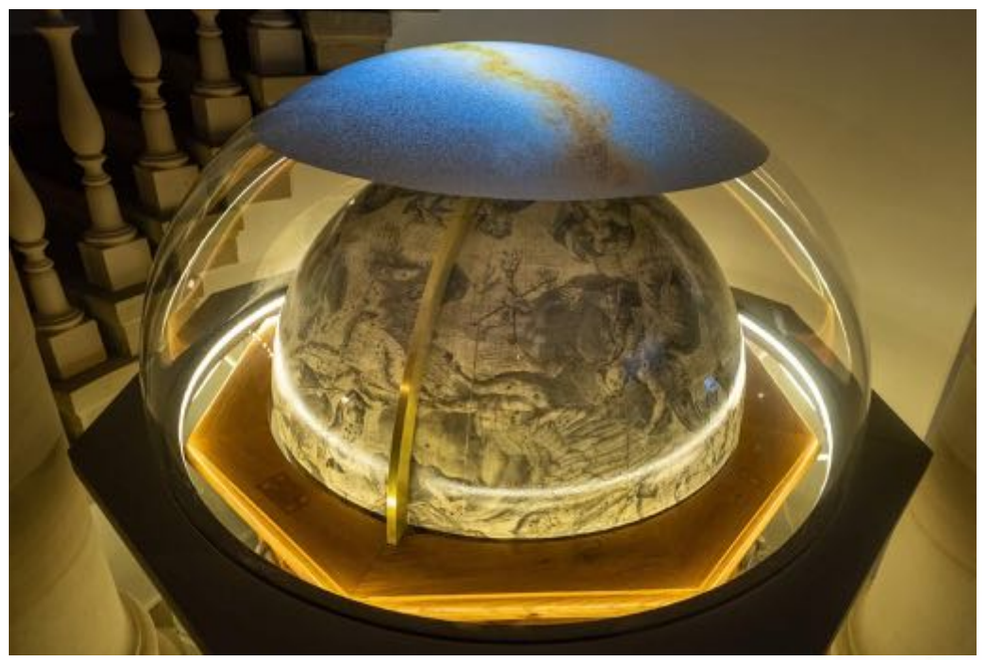
Highlights

- **The VizieR reached a milestone of 20000 catalogues!**

- **Aladin Desktop version 11 - released in April 2020**
  - Improved Time support
  - Advanced data discovery ‘tree’
  - Improved FITS and IVOA features
  - Grav. Wave HEALPix format compatibility



- **‘Local’ highlight**
  - **Aladin Lite on the Coronelli globe**
  - HiPS version of the scanned globe
  - CDS data in interactive display

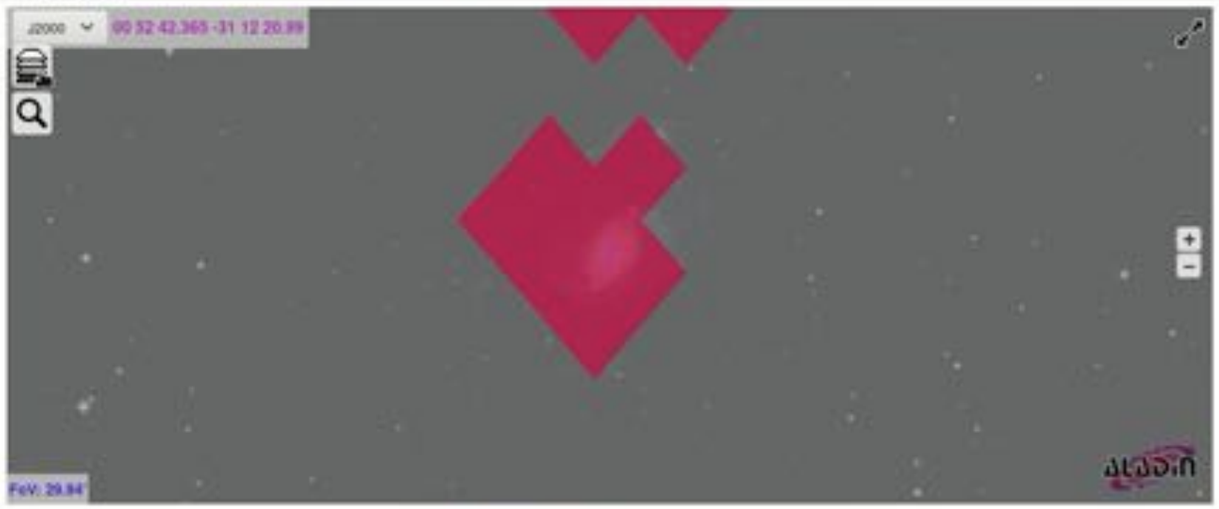


CDS in the community

CDS @ eas Leiden 2020 Virtual

- **CDS Lunch Session** at the European Astronomical Society conference
- ~85 participants
- Presentations: CDS, SIMBAD, Vizier, CDS by Python
- Community feedback and questions

```
In [2]: moc_1 = cds.find_datasets(meta_data="ID=CDS/P/GALEXGR6/AIS/FUV", return_moc=True)
In [3]: moc_2 = cds.find_datasets(meta_data="ID=*HST*", return_moc=True)
In [4]: moc_3 = moc_1.intersection(moc_2)
In [5]: aladin= ipyal.Aladin(survey='P/DSS2/color', target='NGC289', fov=0.5)
        aladin
```




```
In [6]: aladin.add_moc_from_dict(moc_1.serialize(format='json'),
                                {'color': 'grey', 'opacity': 0.7, 'name': 'GALEX'})
        aladin.add_moc_from_dict(moc_3.serialize(format='json'),
                                {'color': 'crimson', 'opacity': 0.7, 'name': 'Intersection'})
```

CDS in the community

ESA/ESO SCIOPS multi-messenger meeting, ESAC Madrid, Spain, 19-22 November 2019
Addressing common challenges for FAIR data in astronomy (M. Allen).

The 5th Scientific Writing for Young Astronomers, Kunming, China, 6-10 January 2020
Lectures/tutorials on *From your paper to VizieR and SIMBAD* (L. Cambrésy).

Kavli-IAU Workshop: International co-ordination of multi-messenger transient observations in the 2020s and beyond. Cape Town, South Africa, 3-7 February 2020
Role of the IVOA presentation (A. Nebot).

Workshop for the INSU prospective challenge 14 *"Open Access to Scientific Data"* / Atelier prospective INSU défi 14, "Accès Ouvert aux Données Scientifiques", 20-21 January 2020
Contributions to INSU prospective from CDS members. C. Bot co-led the workshop and led the writing of the conclusions.

ADASS - Groningen Oct 2019, Virtual Nov 2020 - major event for CDS presenting results and progress to the astronomy scientific and data/software community

(**LISA** - Library and Information Science in Astronomy - postponed to June 2021)

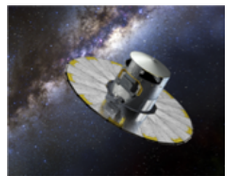
CDS mission

- **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

MONTHLY NOTICES of the Royal Astronomical Society

+ …

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

EUROPLANET 2024 Research Infrastructure

Data e-Infrastructures

RDA
RESEARCH DATA ALLIANCE
+ RDA Europe

CORE TRUST SEAL

WORLD DATA SYSTEM

Renewed CDS membership of WDS

- The World Data System
  - Interdisciplinary Body of the International Science Council (ISC; formerly ICSU)
- Trusted Scientific Data Services and Data Communities
  - '**communities of excellence**' for scientific data services

…goes hand-in-hand with Core Trust Seal certification

Staff

Permanent staff:

11 - Researchers
(8 CNAP, 3 CNRS)
8 - Software engineers
11 - Documentalists

Contract staff:

3 - Engineers
2 - Documentalists
2 - Postdoc researchers

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Direction and Administration

M. Allen (Director)

P. Fernique (Technical Lead), L. Arbousse, C. Halter

Science

M. Allen
C. Bot
L. Cambrésy
S. Derriere
F. Genova
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
A. Oberto
F-X. Pineau
A. Schaaff
G. Mantelet

Documentation

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

Support (shared with Observatoire de Strasbourg)

S. Langenbacher, V. Trimbour, C. Saillard, T. Keller

Post-doctoral Researchers
Y. Stein
K. Lutz

Project support
M. Baumann
H. Heinl

Ph.D Students
T. Lizee
T. Roland

Software Development Interns
~10-12 interns, short term contracts

Staff - recent and coming changes

Permanent staff:

— CNRS competition for documentalist: Esther Collas successful!

— Transition to Emeritus Status: Francoise Genova - December 2020
(Not time to say goodbye yet!!)

— Return from 1 year leave (September 2020): Michaela Buga

Continuing UNISTRA contract: Grégory Mantelet CDI - starting Jan 2021

Contract staff changes:

— Recruited ESCAPE technical support engineer: Hendrik Heinl (started May 2020)
— ESCAPE engineer: Matthieu Baumann (current contract to March 2021)
— Postdoc Yelena Stein - finishing January 2021
— Postdoc Katharina Lutz - finishing March 2021
— Postdoc ESCAPE science support - Stefania Amodeo - starting March 2021
— Seeking a contract engineer for VizieR in 2021
— Documentalist contracts to be defined in 2021

CDS Strategy

- **Constantly evolving** — based on scientific needs
- In 2020 — focused on responding to immediate issues
- **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**

- Strategy / plan to be prepared for HCERES exercise 2021/22 for 2023-27 period

National and European Landscape

Elements that define high level policies:

- **French National Roadmap for Research Infrastructures**
- MESRI National Plan for Open Science
- 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)

Coming soon…

- ASTRONET Science Vision and Infrastructure Roadmap (in preparation now)

Policies are increasingly related to FAIR principles and Open Science

FAIR

- **F**indable, **A**ccessible, **I**nteroperable, **R**eusable

Open Science

- Data sharing with open and seamless services to analyse and reuse research data to improve science

Stewardship

- Human skills for curation, quality content, data management, services

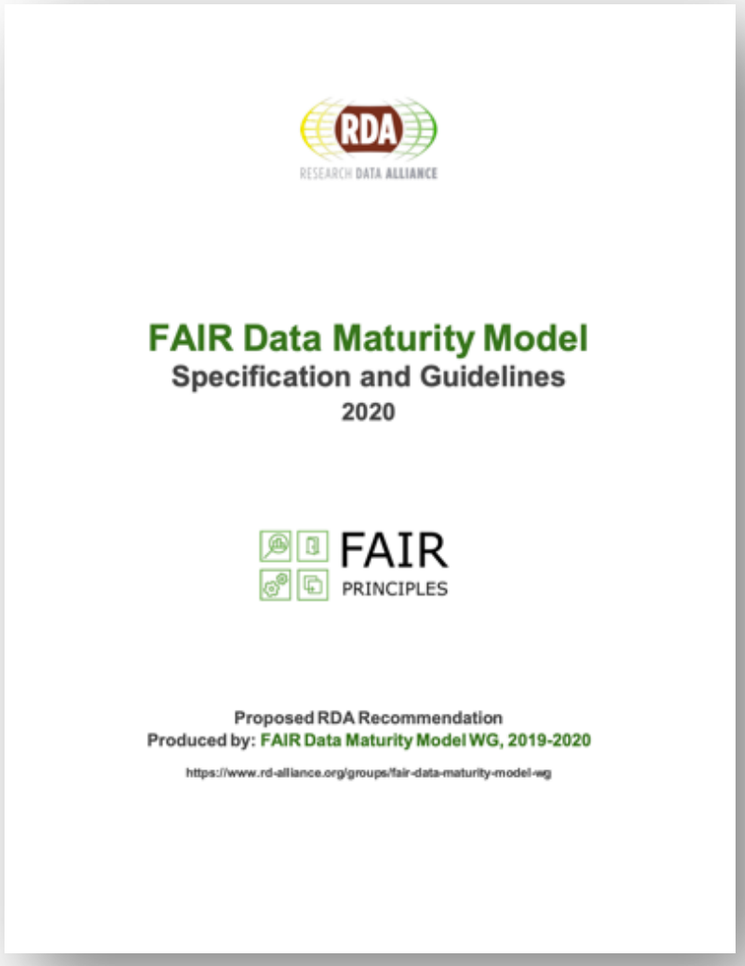
FAIR principles for Open Science are prominent in the current ”policy landscape”

- National Plan for Open Science *2018*
- National Infrastructure Roadmap *2018*
- ESFRI Roadmap *2018*
- (ASTRONET Infrastructure Roadmap *2014*
- EOSC Partnership Proposal, *May 2020*
- EC Strategic Research and Innovation Agenda for EOSC – *v0.8, Oct 2020*

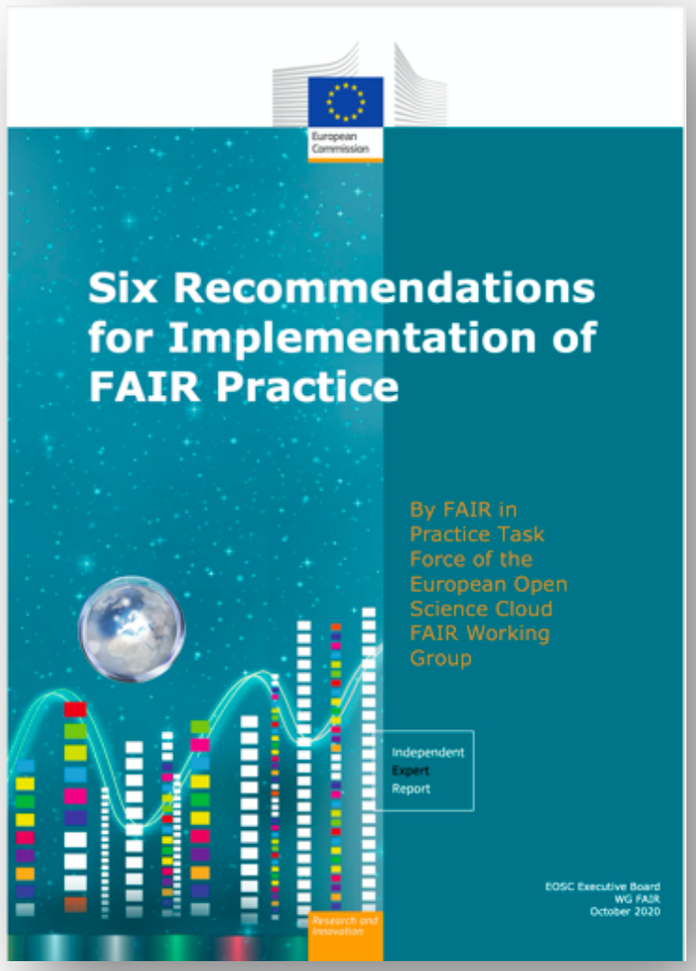
Also locally @ UNISTRA with emphasis on Open Science and Open Access

Specific literature on FAIR

doi: 10.2777/1524



doi: 10.15497/rda00050



doi: 10.2777/986252

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

- 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
  - *See the document of IVOA contributions*
  - *See the demo at the end of the day!*

Projects

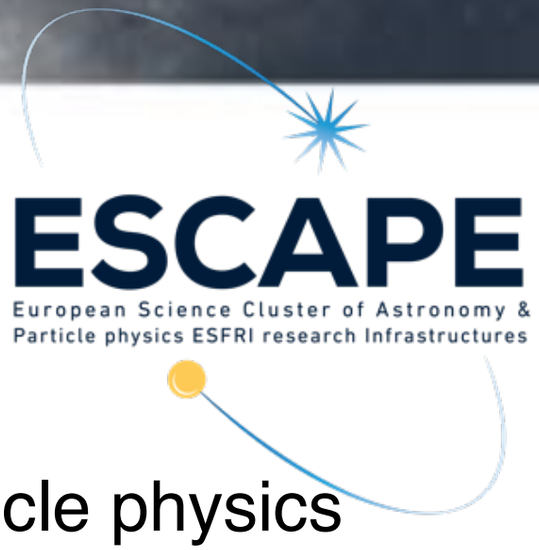
- **ESCAPE -** big project in progress… see next slides…

- **Europlanet 2024** Research Infrastructure (EPN-2024-RI)
  - Small but important CDS participation (~20PM)
- **AENEAS** - finished in December 2019
- **RDA Europe 4** — concluding September 2020

- **XMM2ATHENA** **(***starting 2021)*
  - A. Nebot science, local coordination by ObAS GALHECOS team
  - ~5PM participation of CDS related to X-matching

- **EOSC Future** - large (~40 M€) project of science and e-Infrastructures
  - Passed 1st step of approval, result hopefully known soon
  - 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
- Mid-term review last week (27 November) - *went well!*

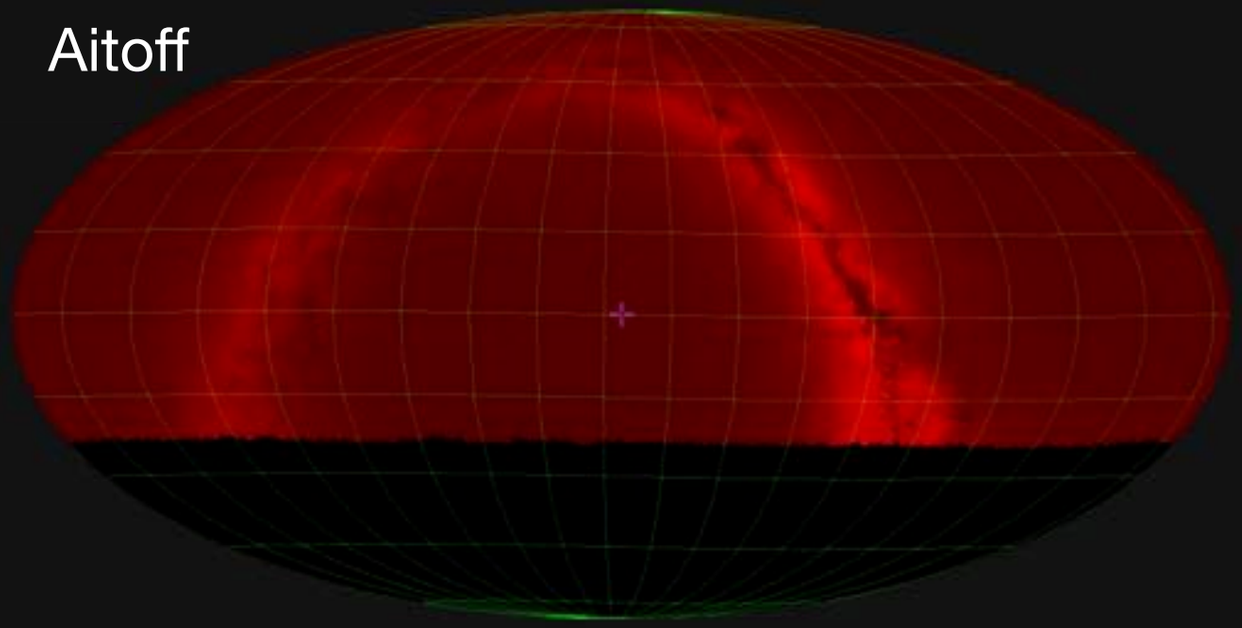
https://projectescape.eu

1.1 Virtual Observatory – part of the ESCAPE cell

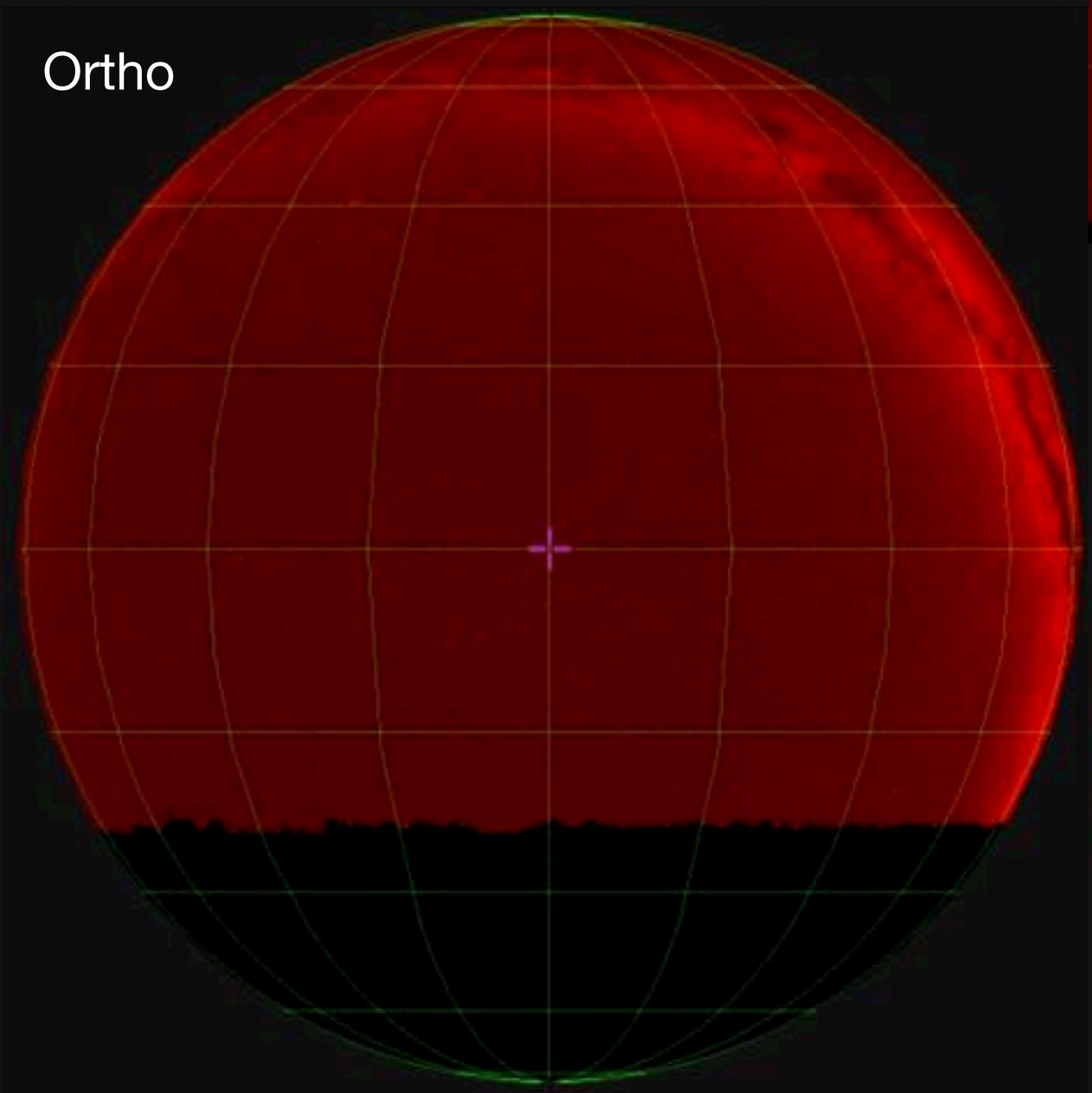
- Connect ESFRI and RI data to EOSC by VO
- Metadata standards based on ESFRI needs
- **Software connections on *deep learning* with WP3**
- **VO connected to storage and computing with WP2**
- **VO data via platform with WP5**

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

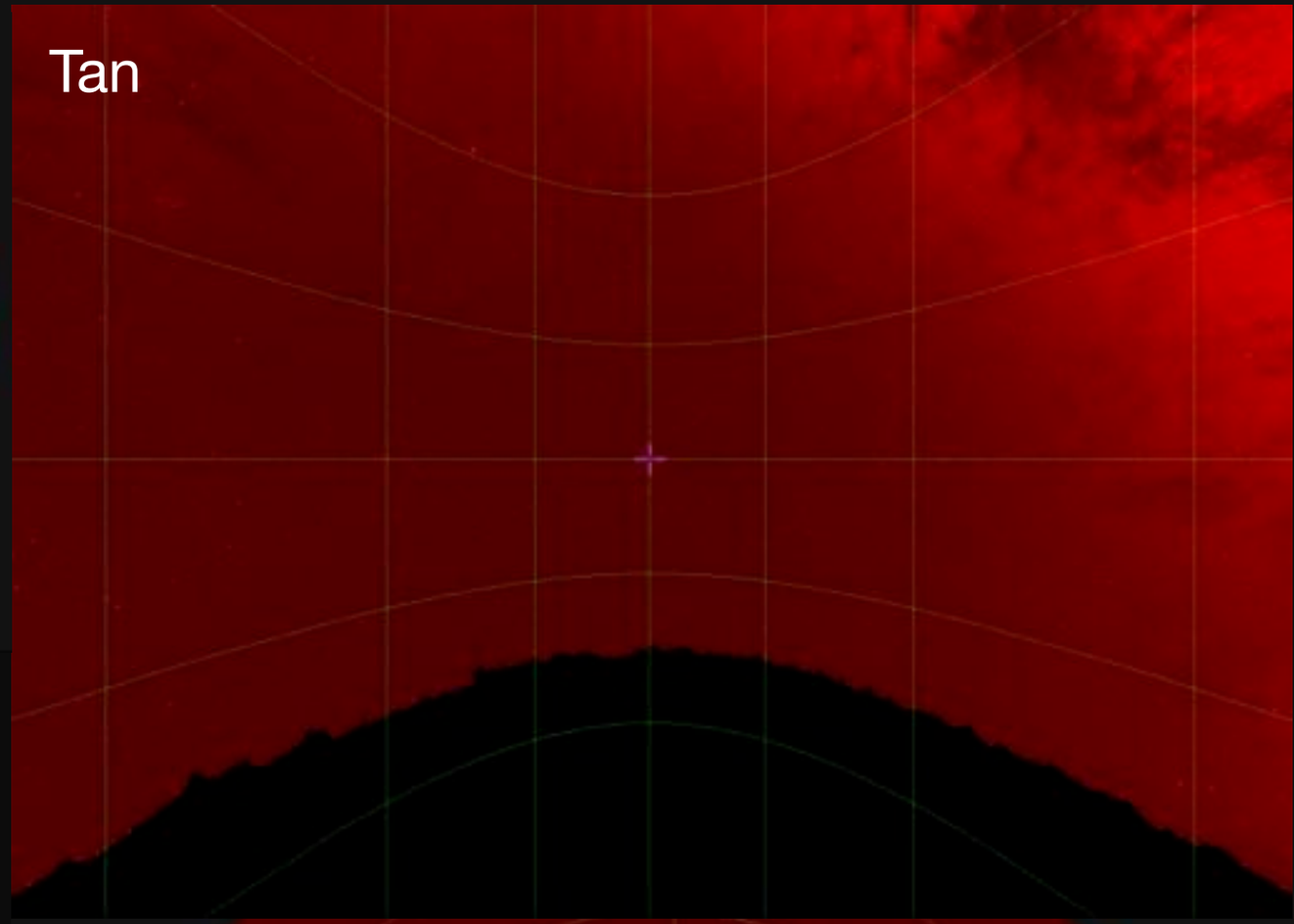
Aitoff



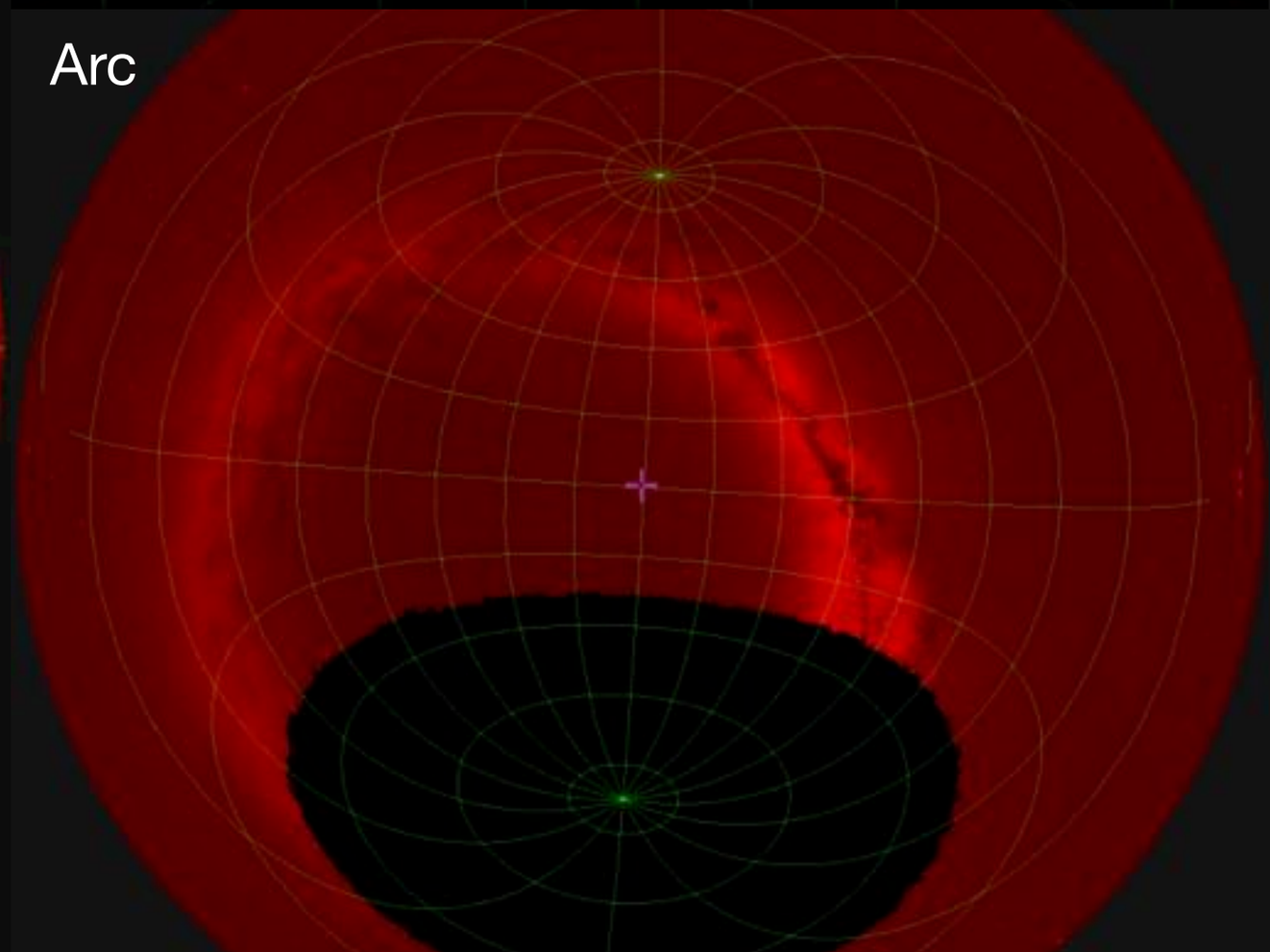
Ortho



Tan



Arc

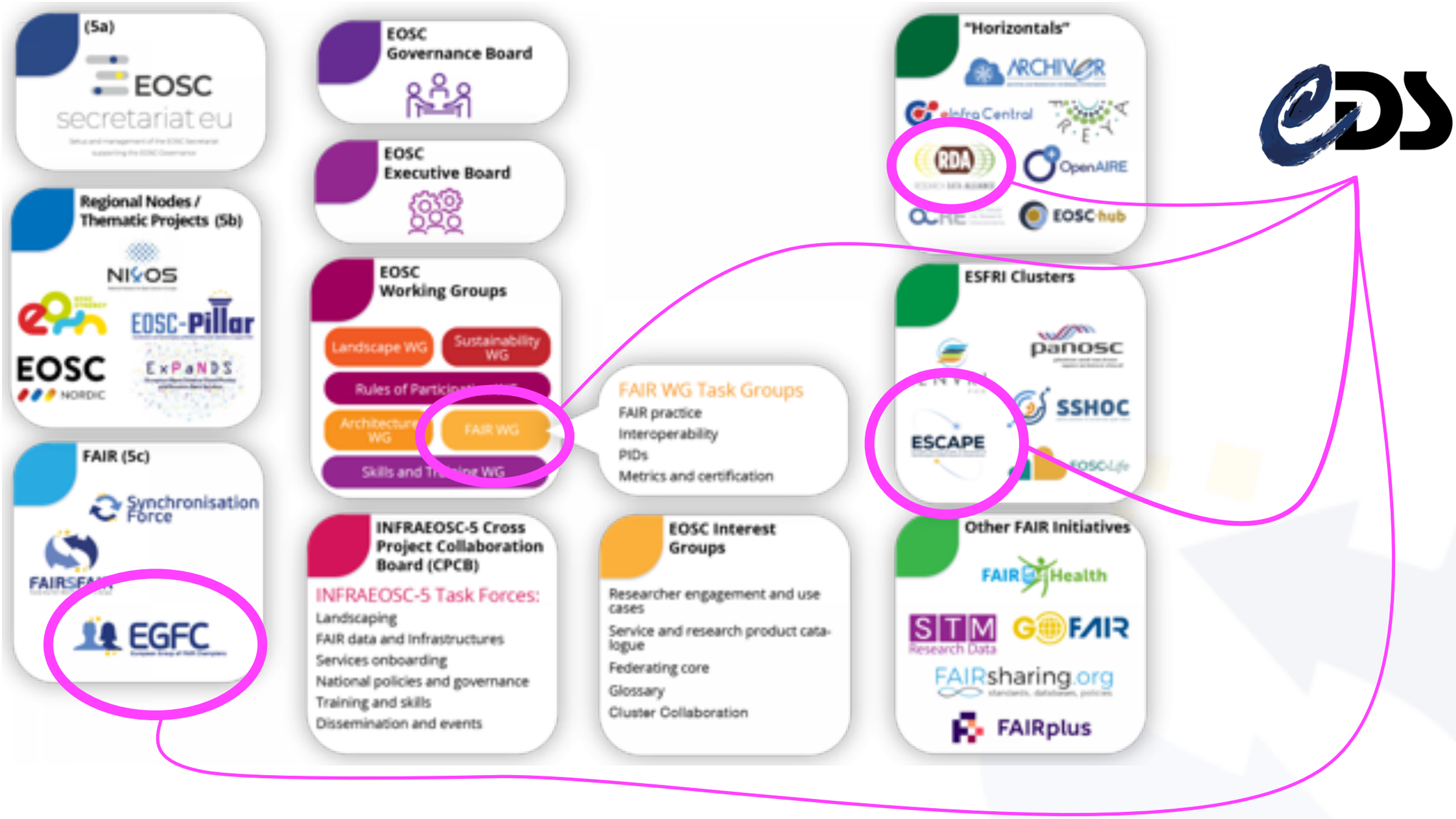


Apogee catalog upon VISTA image survey J, Y, Z bands

1. 1-channel grayscale texture containing the kernels

2. Use of a colormap on the grayscale values

Primary stakeholders in ecosystem



Summary

- A very busy year for CDS with many (continuing) challenges
- Operational stability and continuity throughout
- 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

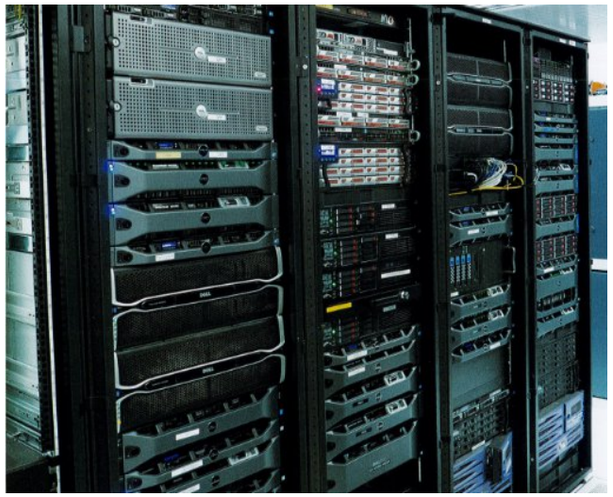
The CDS Information System

Overerview & stats 2020

CDS council – 30 nov & 1 dec 2020

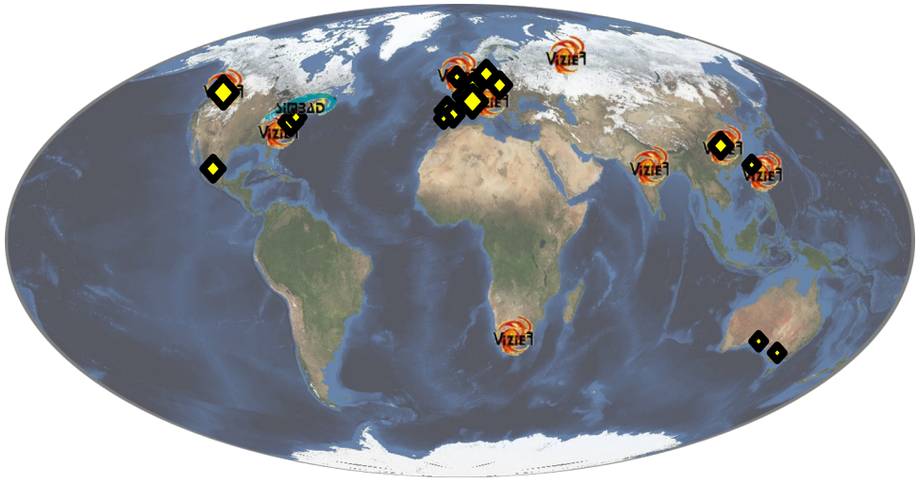
Pierre Fernique
on behalf of all the CDS staff & more

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG



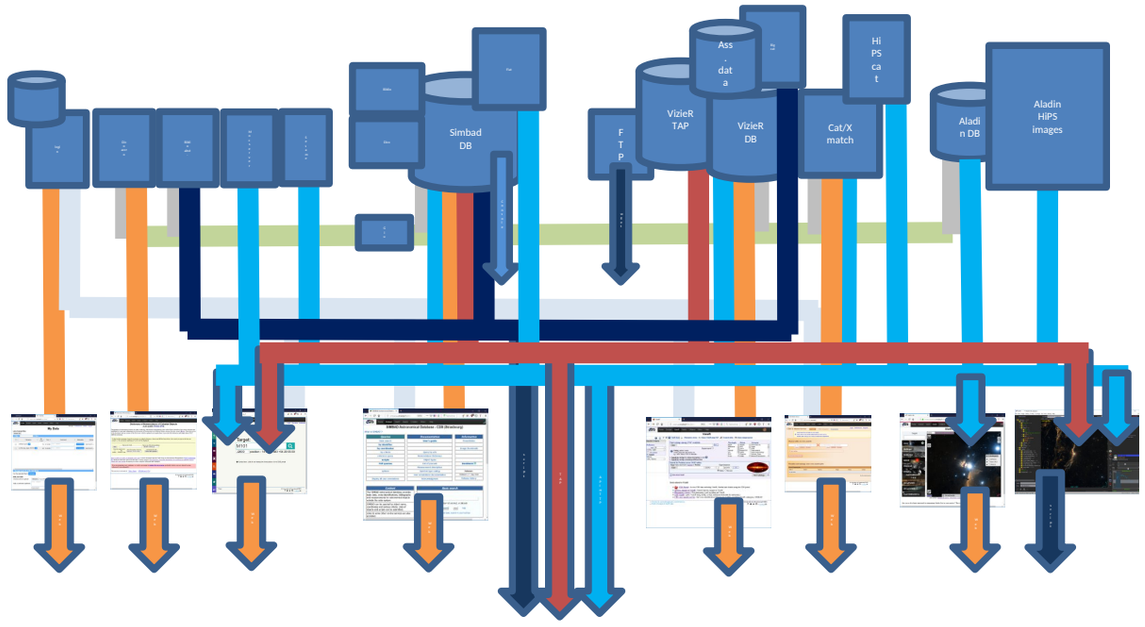
- **12 local machines** (real or virtual) for implementing the 20 sub-components
- CDS **allsky storage system**
  => 1.6PB replicated
- **2 computer rooms**
  => One at the Observatory
  => Replicated elements will move to Unistra Data Center

New
- In test phase
- CDS migration planned for spring 2021



- **10 external sites** (9 VizieR + 1 Simbad mirrors)
- **19 partner HiPS sites** (Aladin tiles)

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



Objects
from litterature

11.5M objects *+9%*

24GB

736K queries/day *+61%*

>177K IP/month *+31%*

Catalogues
from litterature & surveys

20.2K catalogues *+6%*

46TB

520K queries/day *-26%*

>33K IP/month *+7%*

Images
from surveys

869 HiPS *+44%*

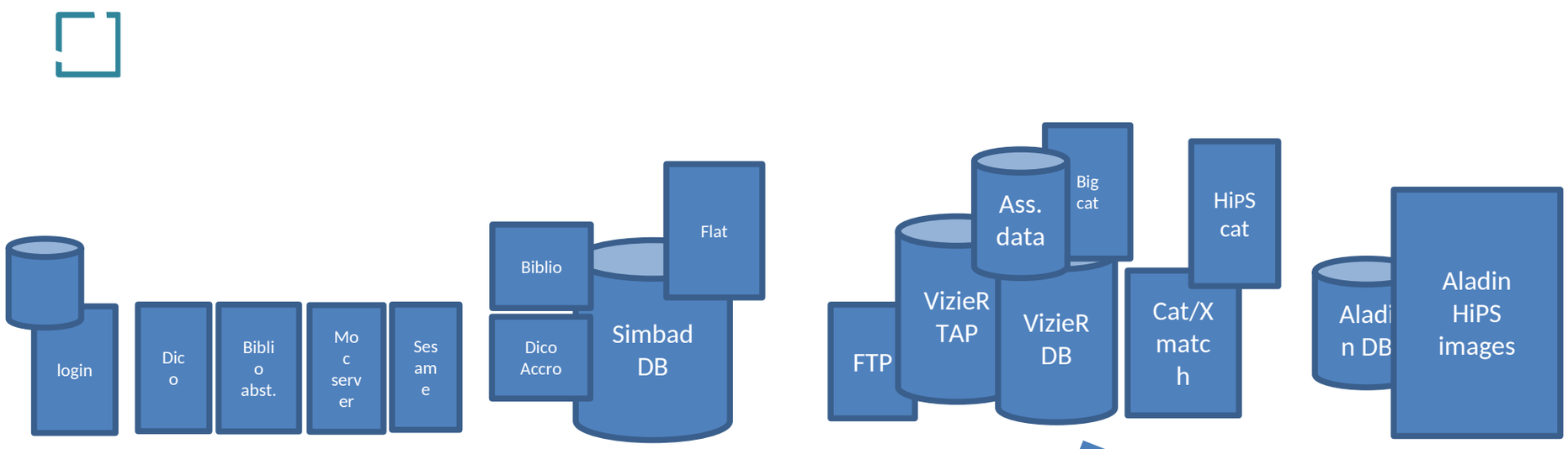
370TB *+41%*

603K queries/day *-29%*

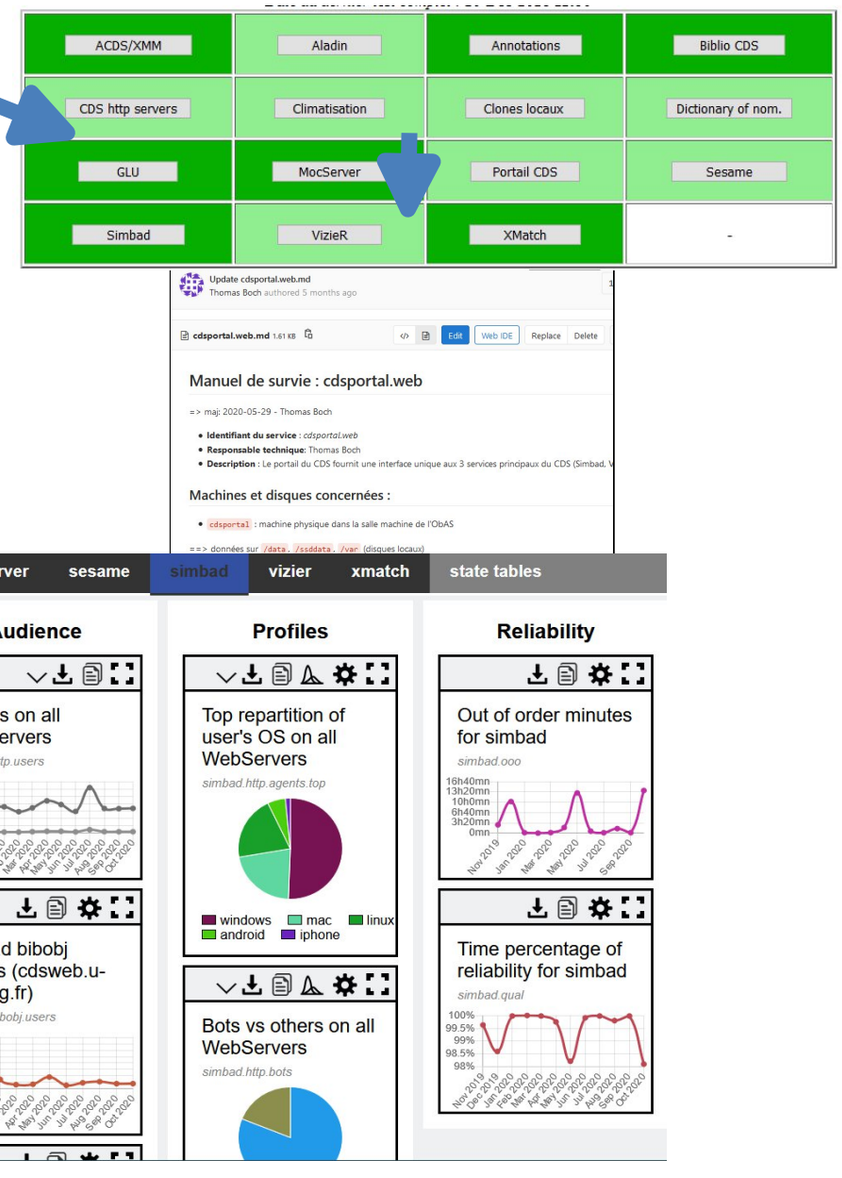
>82K IP/month *+22%*

Sesame	**Xmatch**	**MOCserver**	**CDSlogin**	**Dictionary**
Name resolver	catalog crossmatcher	Resource yellow page	User account & annotations	Nomenclature
~304K queries/day	3K queries/day *+200%*	57K queries/day *+29%*	~500 queries/day	~700 queries/day

- Evolutions:
  - **New internal CDS bibliographical service**
    => support for new DJIN & VizieR table worflows (in deployement)
  - **New Sesame** (prototype phase)
  - **Simbad Web interface** (in progress)
  - **Xmatch code rewriting** (in progress)
    => better perf, better meta data management
- Rationalizations:
  - Flat simbad => will be directly supported by Simbad (in progress)



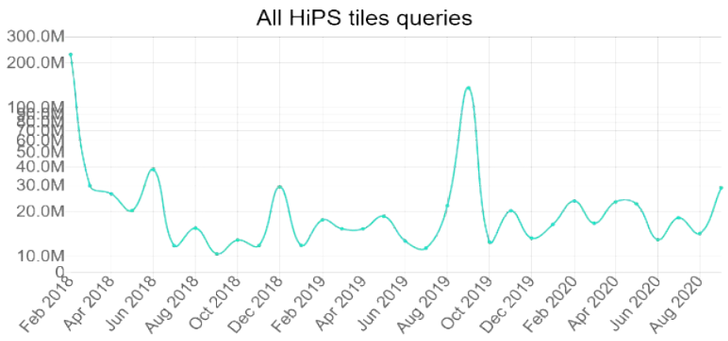
1. **Supervision** & alerts
   => GLU Supervisor
2. **Recovery procedures**
   New => Fully rewrote in 2020
3. **CDS dash board**
   New => Operational

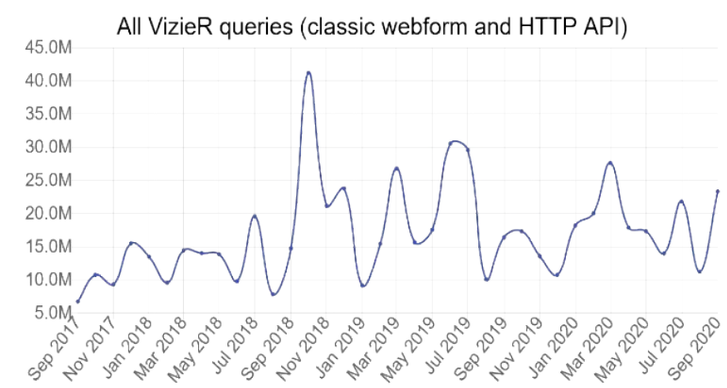


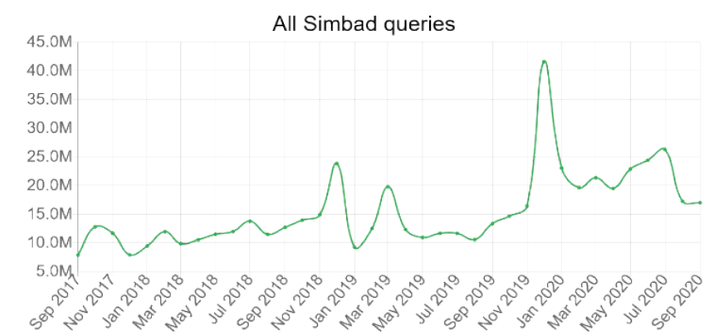
- The **growth of our services continues to be strong**, both in terms of audience and activity
  - almost **2 millions** requests per day
- The **quality of our services**, continues to be high (>**99.7**%)
- Our users have **very diverse profiles**: professional astronomers, amateur astronomers, partner institutes and data centers, general public. i.e. :
  - **80%** of the requests on the CDS are not at the "end of the chain“, but reingested by partners (ADS, ESAC, ESO, STScI, HEASARC, HiPS network, ...) or tools.
  - **60%** of the Simbad audience (IP) is generated from public tools (ex: Stellarium) but for **7%** of the activity
  - The "**Planetarium**" use of HiPS services takes a large part (**34%** Stellarium + <0.5% Digistar)

Two mode activities

=> “continuous use" + "peaks"







Biggest peaks for **Aladin** at a low frequency (1 to 2 times a year)
=> Hips synchronisation

VizieR has to manage a few large "peaks" per semester
=> availability of a new catalog (i.e. Gaia), or specific project or mission that requires VizieR data

Simbad: very high “continuous use" drowns out a good part of the "peaks"

SIMBAD : the bibliographic database

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

CDS Council
December 2020

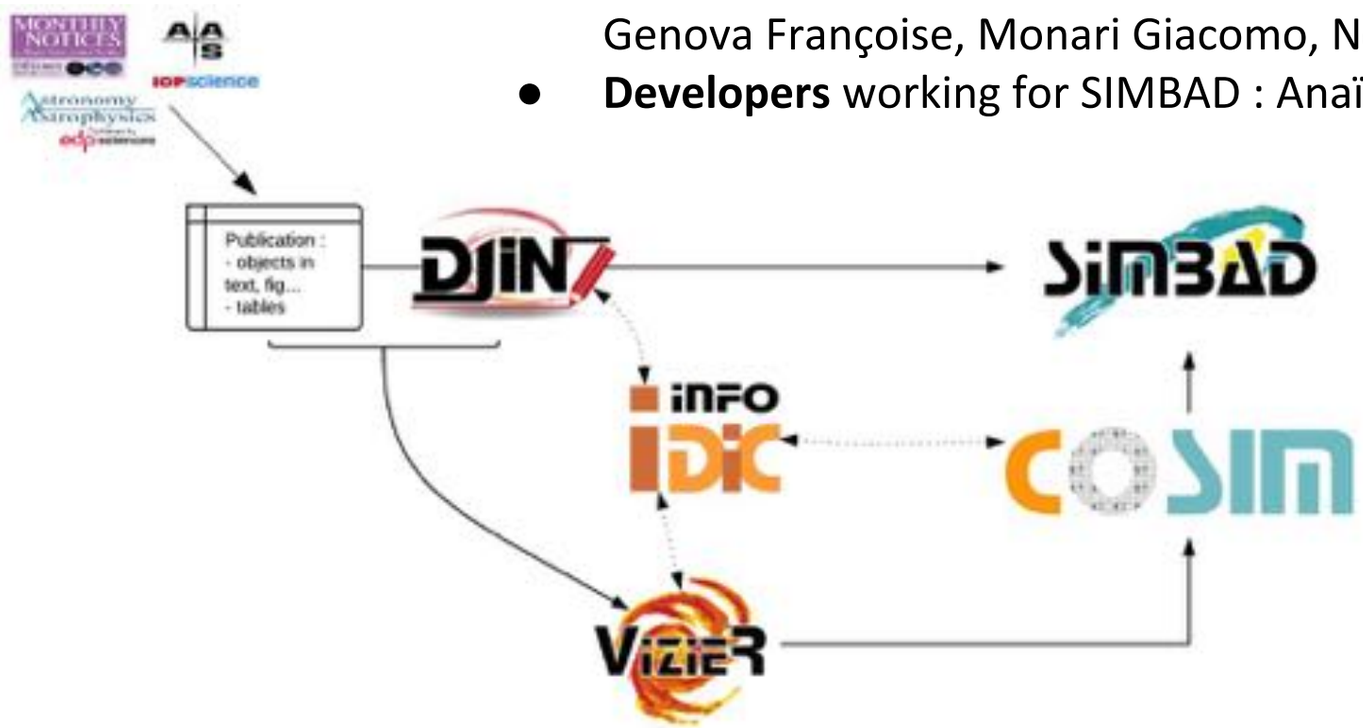
CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

S. Lesteven,
C. Loup,
A. Oberto
B. Vollmer

The Team

- Bibliography & coordination : Soizick Lesteven
- Scientific content : Cécile Loup
- Database & softwares : Anaïs Oberto
- Nomenclature : Bernd Vollmer

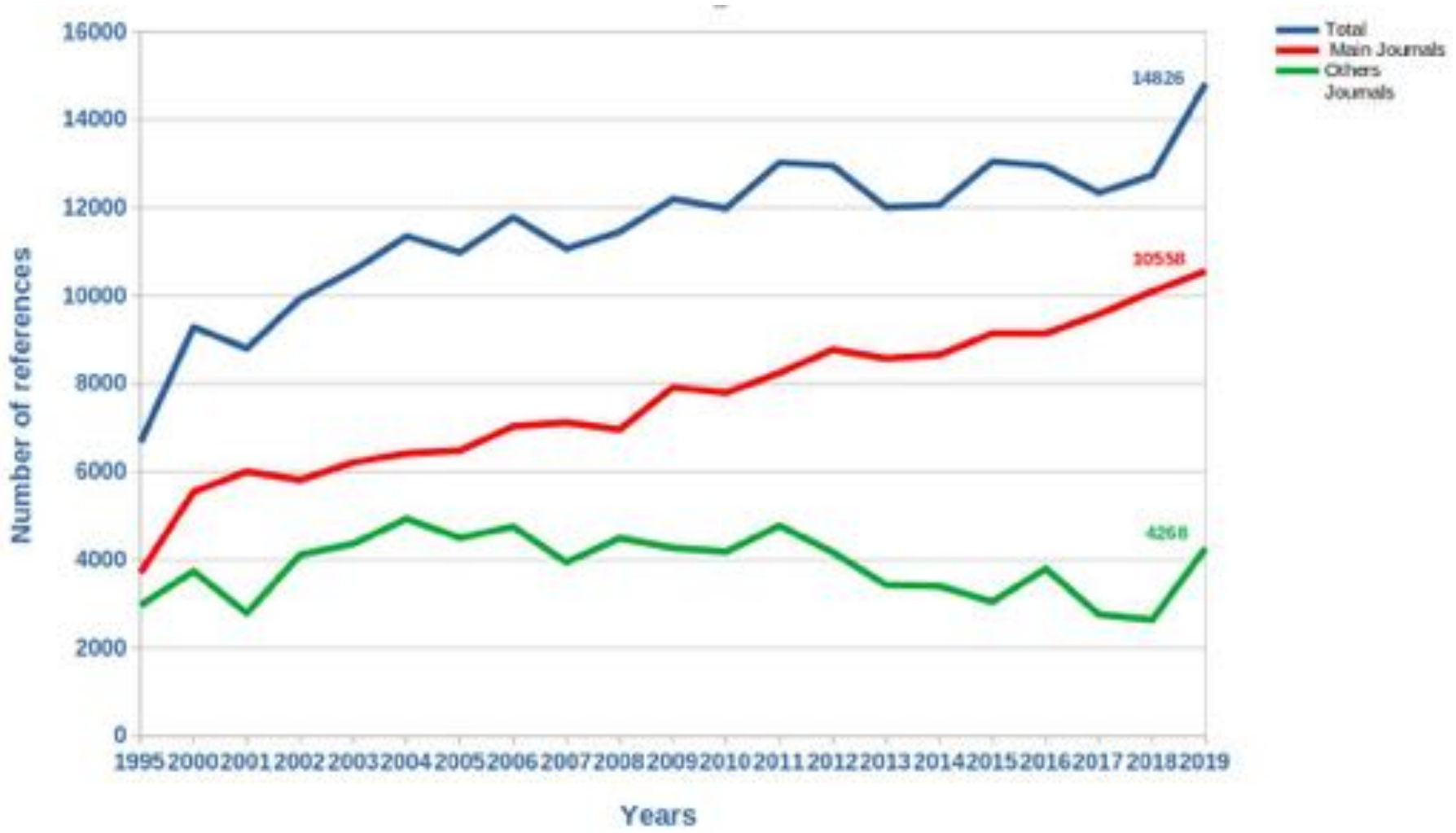
- **Documentalists** (Data Stewards) :
  - Nomenclature (1.5) : Marianne Brouty, Fabienne Marquis
  - Ingestion of references via DJIN (3.5) : Aline Eisele, Evelyne Son, Magali Neuville, Philippe Vonflie
  - Ingestion of lists of objects via COSIM (4) : Catherine Brunet, Esther Collas, Fabienne Marquis, Katia van der Woerd, Mihaela Buga, Emmanuelle Perret
- **Astronomers** involved in scientific content : Bot Caroline, Cambrésy Laurent, Derrière Sébastien, Genova Françoise, Monari Giacomo, Nebot Ada, Ocvirk Pierre, Siebert Arnaud, Vollmer Bernd
- **Developers** working for SIMBAD : Anaïs Oberto, Grégory Mantelet



The Content

- 381,441 References +15,000
- 11,565,529 Objects +587,000
- 36,674,509 Identifiers +1,126,509
- 15,000 Acronyms +300

Number of references processed by SIMBAD by published year 1995 - 2019

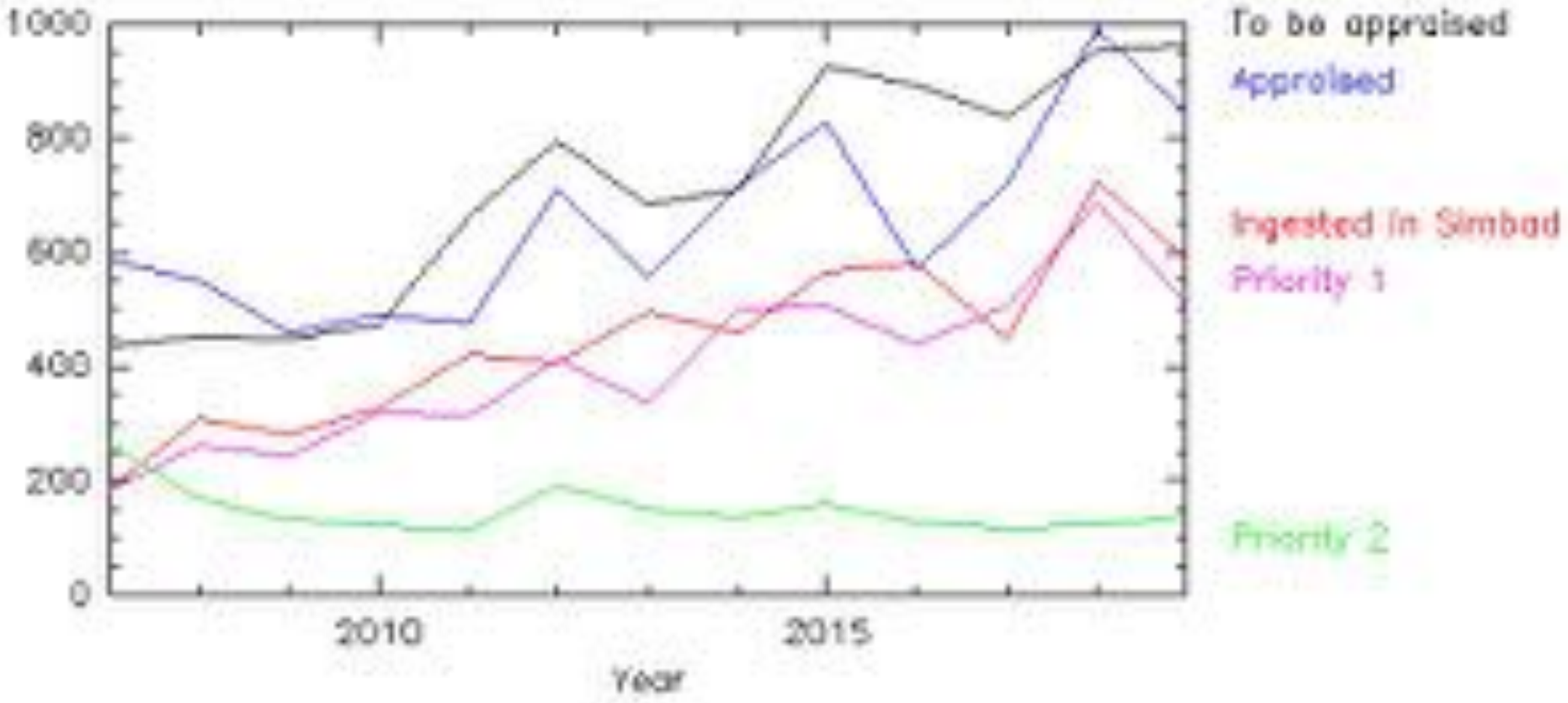


- All efforts are done to keep the quality of the process

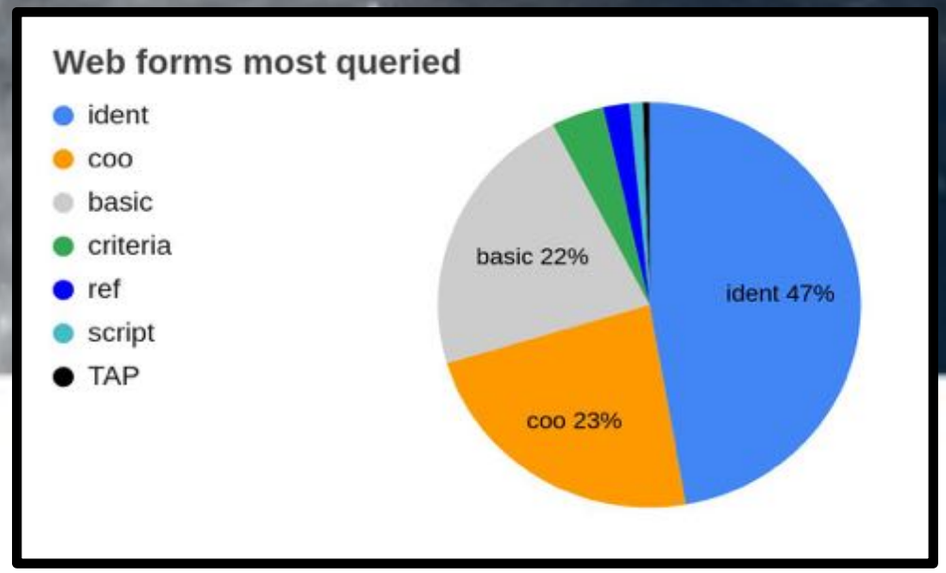
Ingestion of Tables of Objects

Impact of the lockdown : negligible on the ingestion in Simbad, but some backlog on the appraisal.

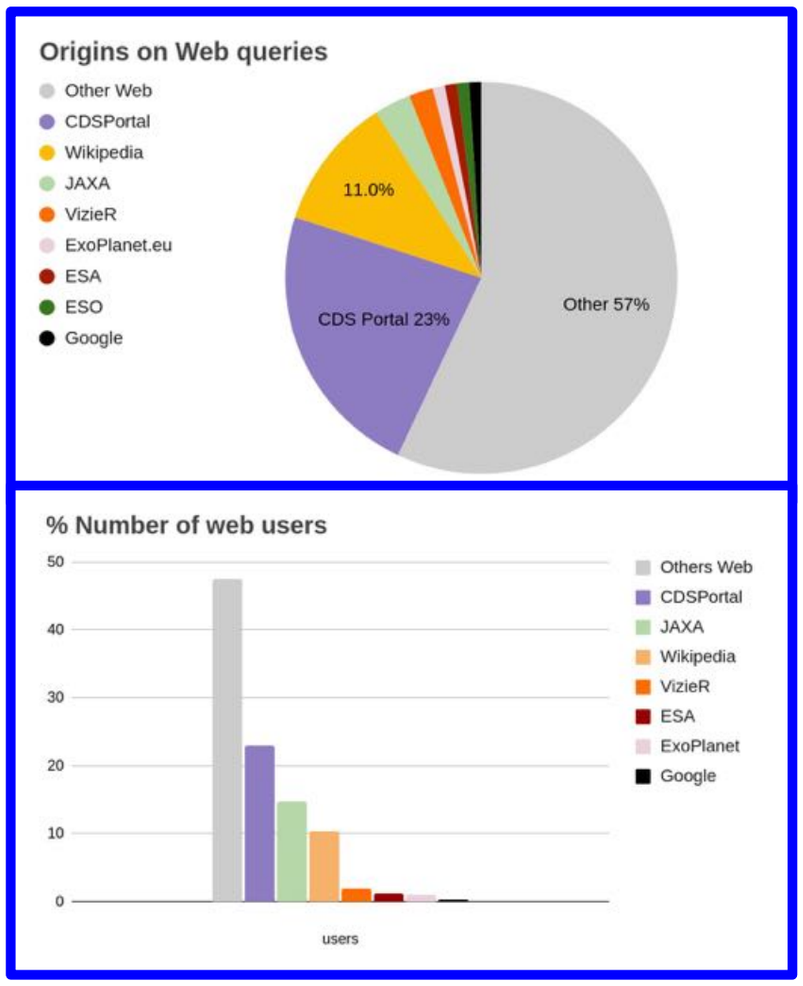
Number of references with table(s) of objects



Usage 2020

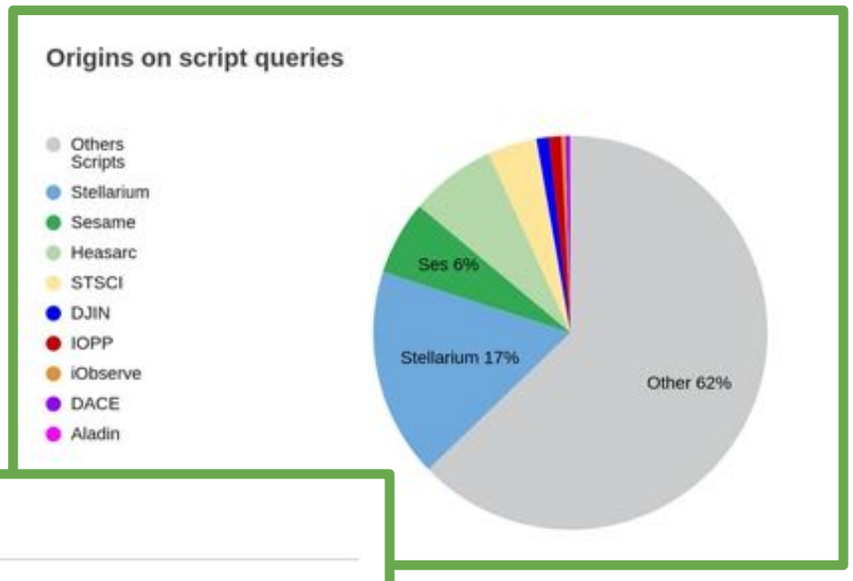


Interactive Web pages : **5%** of usage

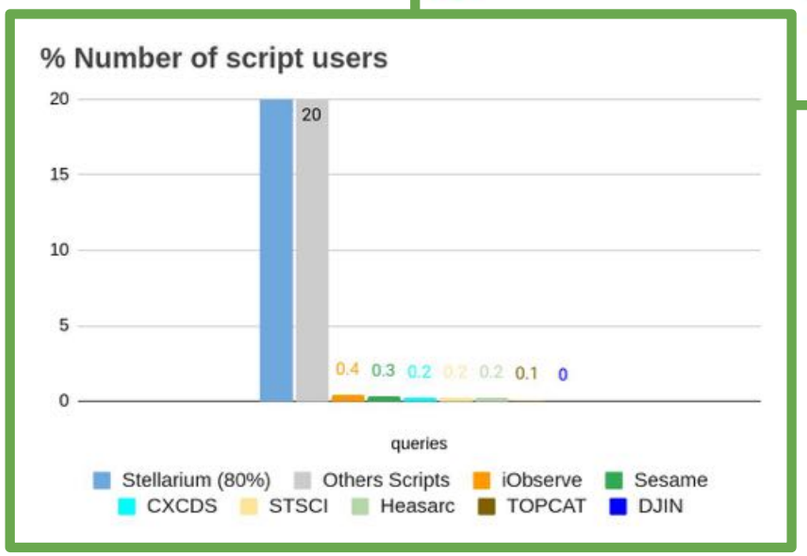


Users: ~ 25K/month

Scripts origins: **95%** of usage (~ 600K/day)
(Others -> Mostly individual Python/TAP)



Users: ~ 160K/month

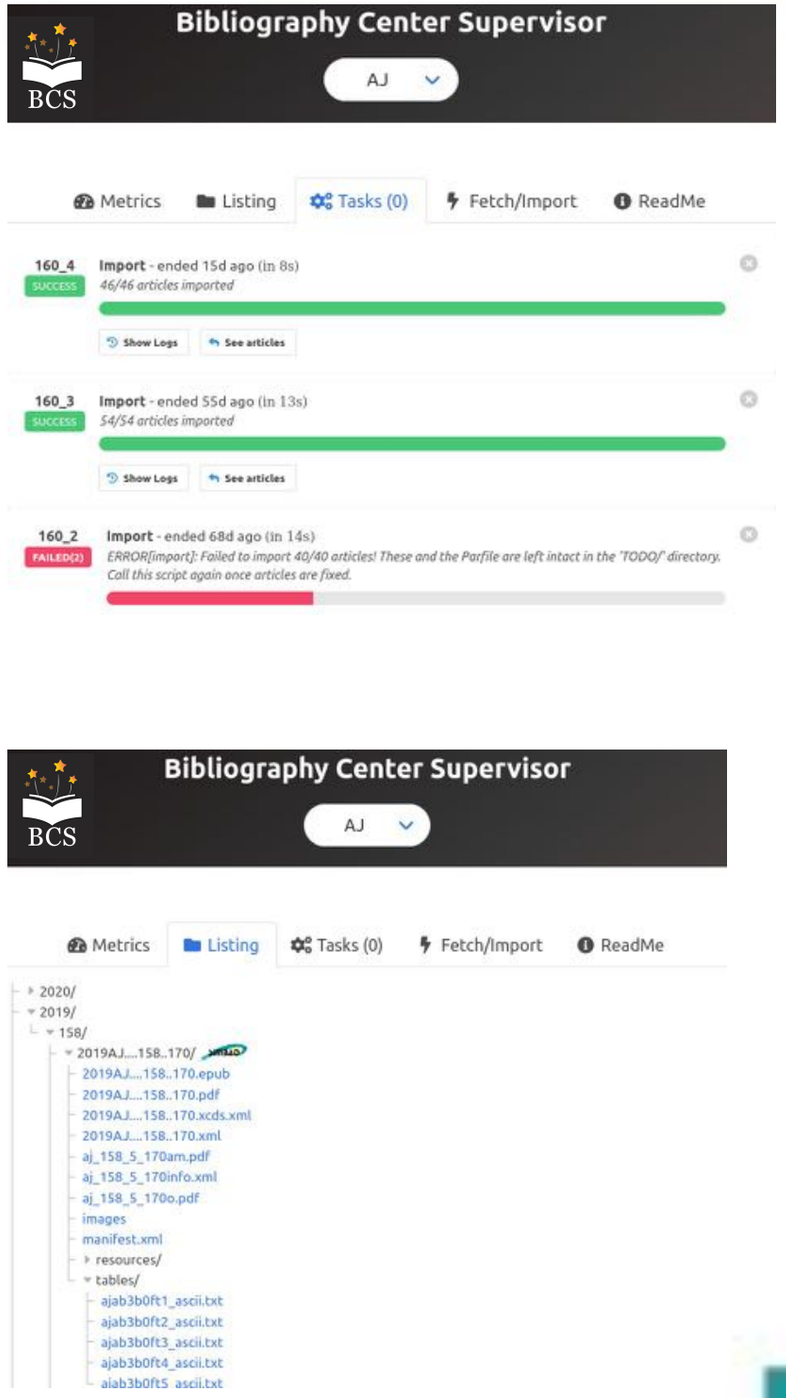


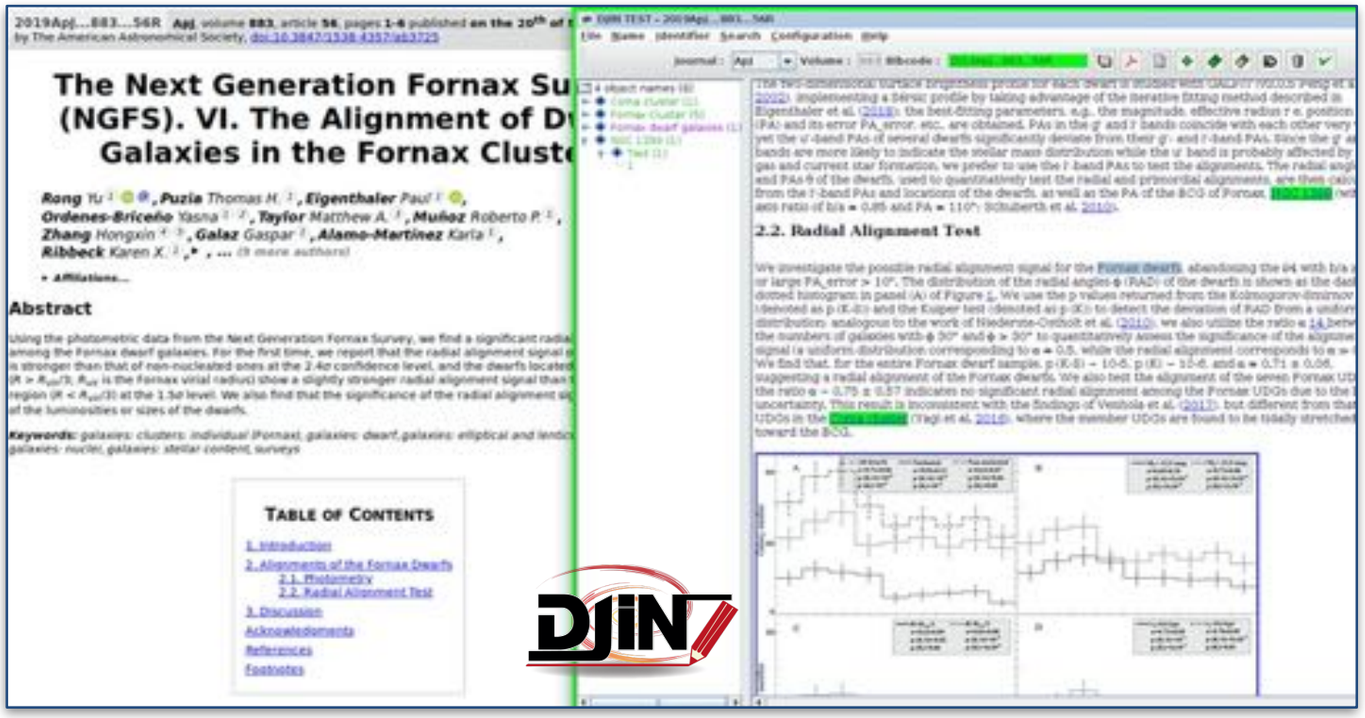
Infrastructure still in progress

- Renew old server machine, and evolving Java version
- Working on a “light” mode avoiding a Java server, but more infrastructure to synchronize clients
- New API to query Simbad in development

DJIN : Internal tool to extract object name in text

- Used currently
- Still some visual features to improve
- Input pipeline to help identifying authors in progress



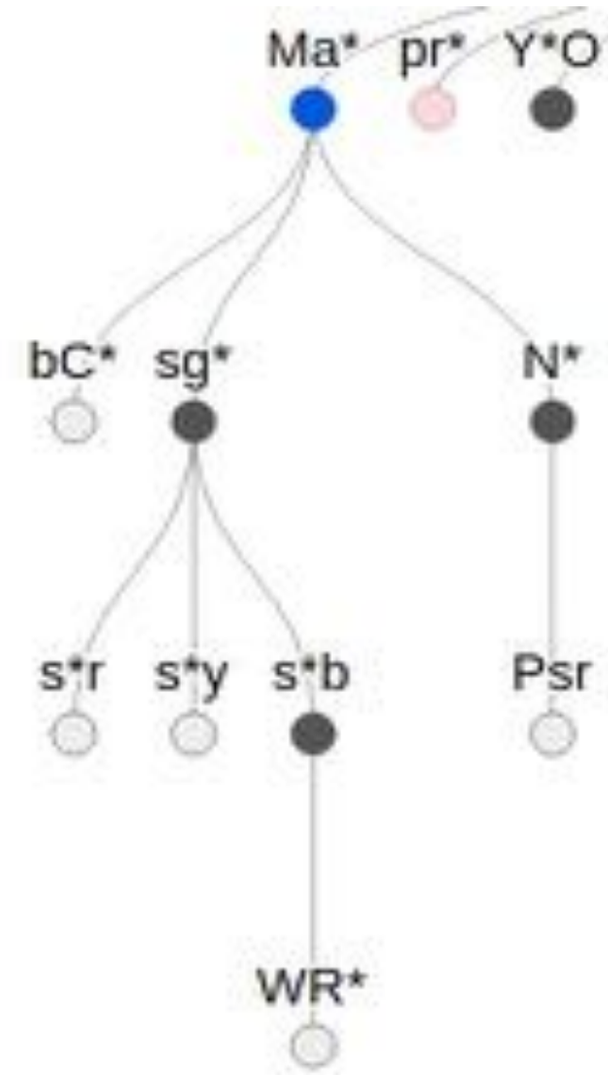


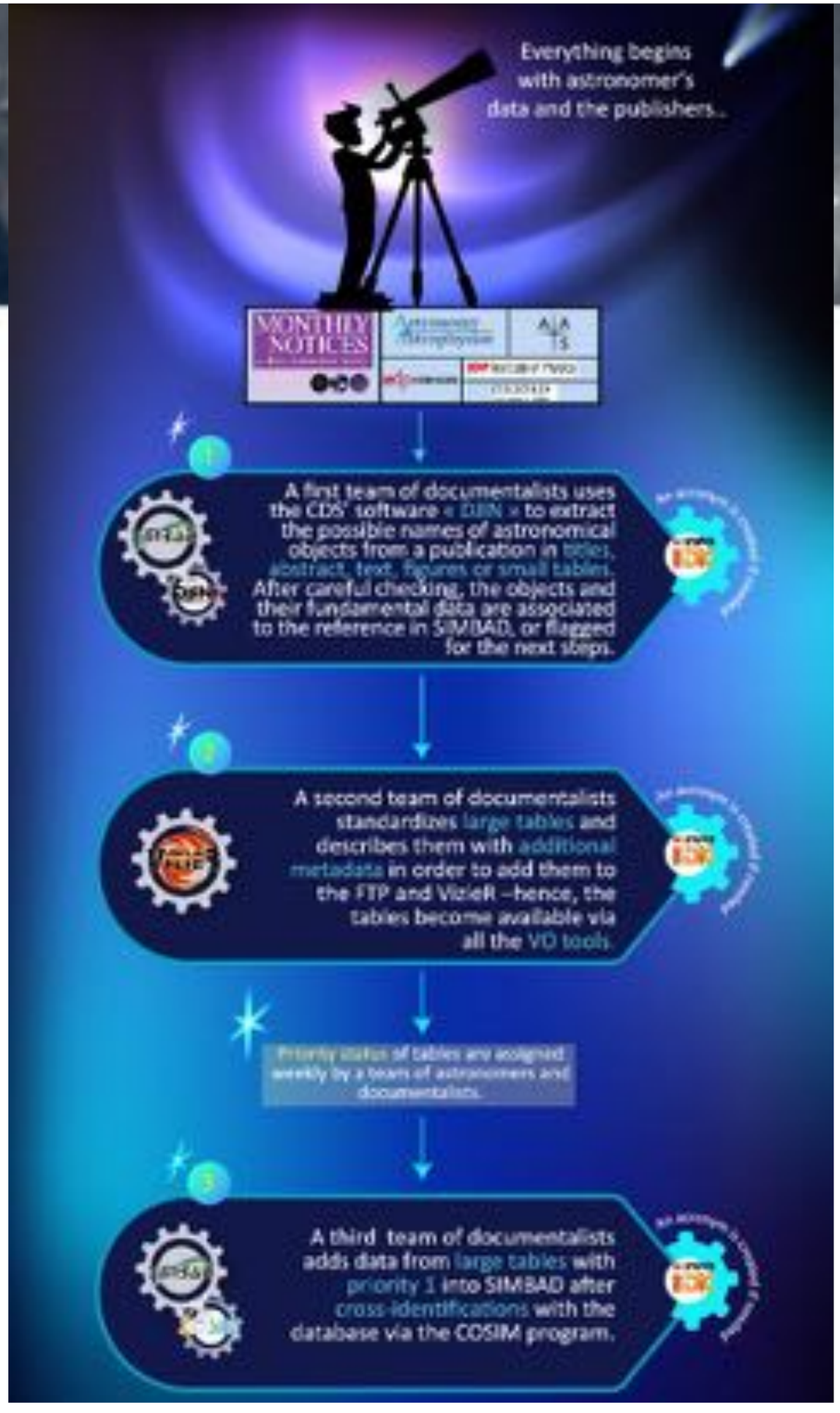
Reorganization of the 230 Object Types

New list, hierarchies, compatibilities : implemented in COSIM, improvement of efficiency and scientific content

1.1. Massive Stars and their Remnants

[Ma*]	[Ma?]	10	Massive Star or its Remnant	Initial mass > 8-10 Mo
bC*	[bC?]	1	beta Cep Variable	
sg*	sg?	4	Evolved Supergiant	Luminosity type 0,Ia,Iab,(I). Includes A-type SG
s*r	s?r	3	Red Supergiant	SpT like K/M 0,Ia,Iab,(I)
s*y	s?y	3	Yellow Supergiant	SpT like F/G 0,Ia,Iab,(I)
s*b	s?b	3	Blue Supergiant	SpT like O/B 0,Ia,Iab,(I)
WR*	WR?	2	Wolf-Rayet	SpT like W
...			Luminous Blue Variable	Maintype = s*b and SpT like LBV
N*	N*?	2	Neutron Star	
Psr		1	Pulsar	




Everything begins with astronomer's data and the publishers..
MONTHLY NOTICES
A first team of documentalists uses the CDS' software « DRN » to extract the possible names of astronomical objects from a publication in titles, abstract, text, figures or small tables. After careful checking, the objects and their fundamental data are associated to the reference in SIMBAD, or flagged for the next steps.
A second team of documentalists standardizes large tables and describes them with additional metadata in order to add them to the FTP and VizieR –hence, the tables become available via all the VO tools.
Priority status of tables are assigned weekly by a team of astronomers and documentalists.
A third team of documentalists adds data from large tables with priority 1 into SIMBAD after cross-identifications with the database via the CCSIM program.

Scientific council 2020

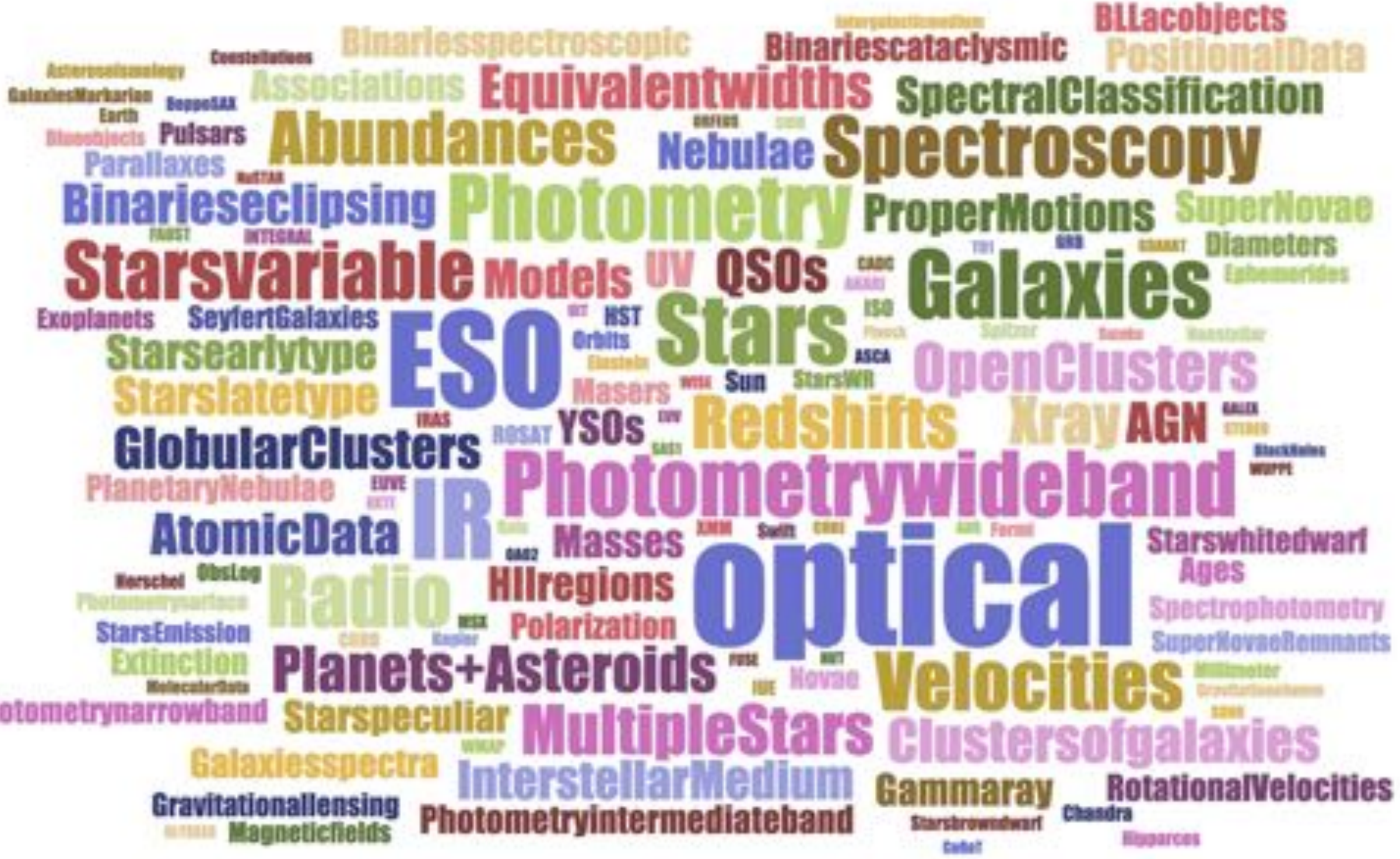
VizieR Staff and contributors:

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

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

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

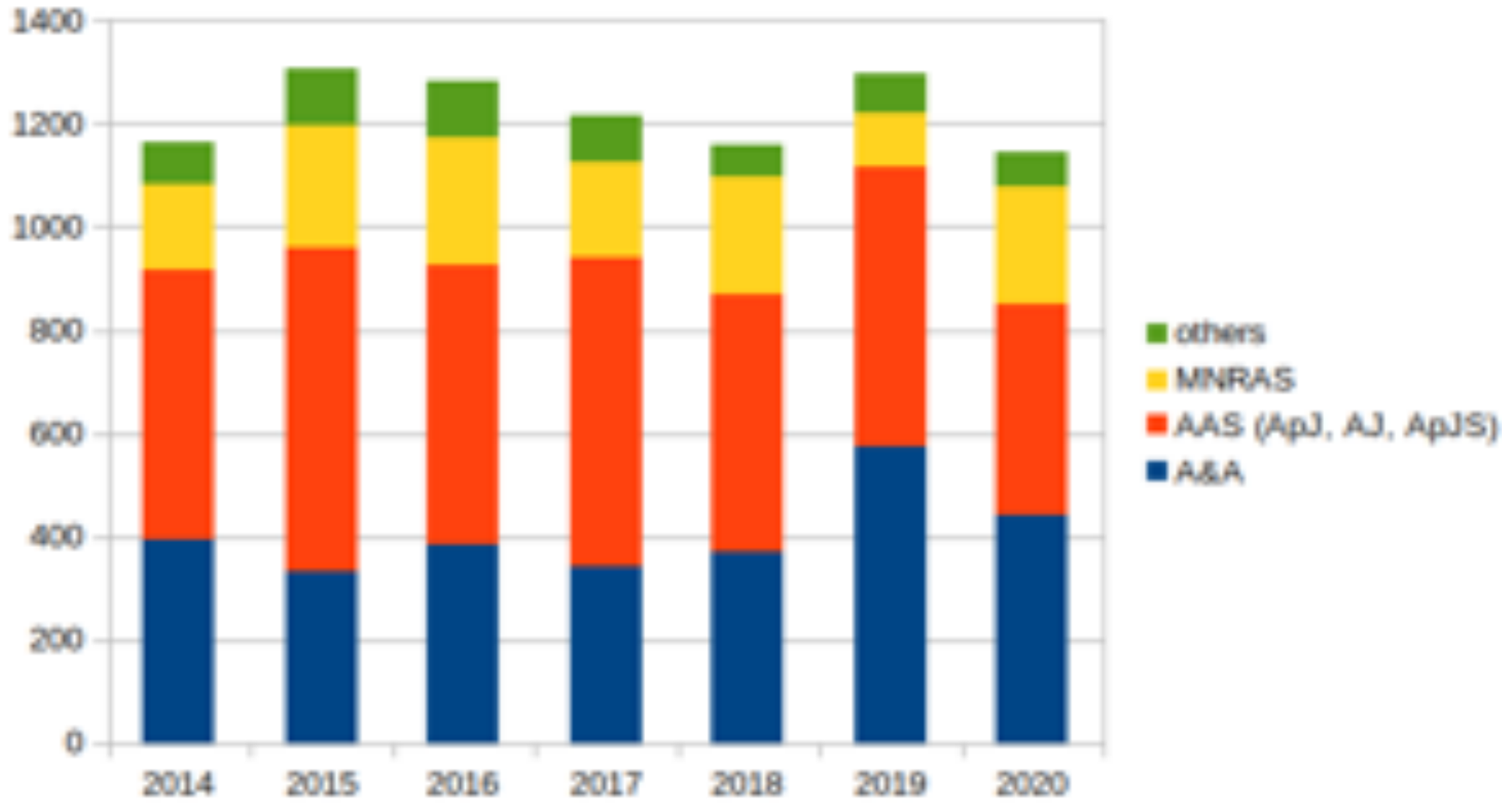
VizieR content - I



VizieR content - II

Ingestion statistitics :

- A steadily increasing volumetry : 20,379 catalogues, 45,399 tables, 33 billion records, 46TB
- Concerted action with A&A to keep volume under control (prioritization)



VizieR content – III : Very large catalogs & co

Very large catalogs:

- Tess Input Catalogue
- Dark Energy Survey DR1
- Gaia DR2 vs_AllWISE YSO
- UNWISE
- IGAPS DR1
- VVV VIRAC PM DR4.1
- Gaia DR2 Extinction
- CatWISE

“Thick” catalogs: > 300 columns

- APOGEE-2 DR16 (342 columns)
- 4FGL (343 columns)

In progress:

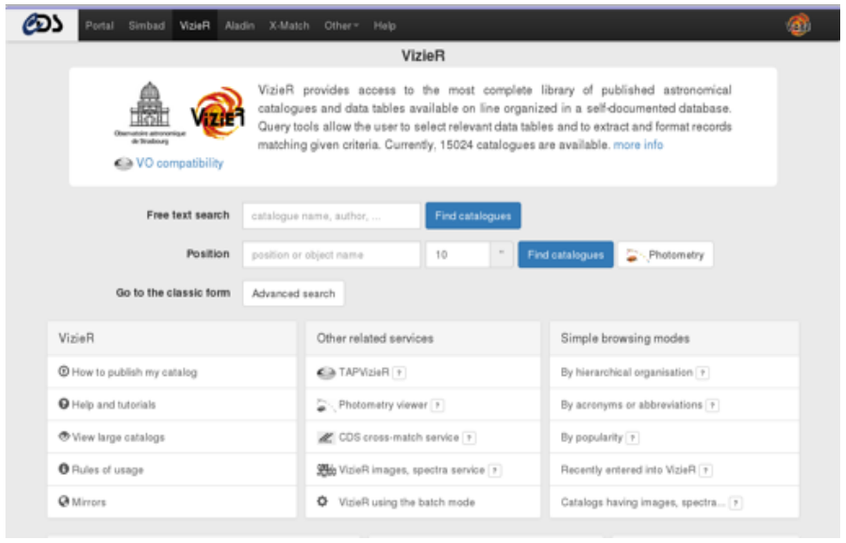
- Gaia EDR3 (Dec. 3rd!!!)
- SDSS DR16 (stdby, some problems)
- Pan-STARRS DR2 (stuck)

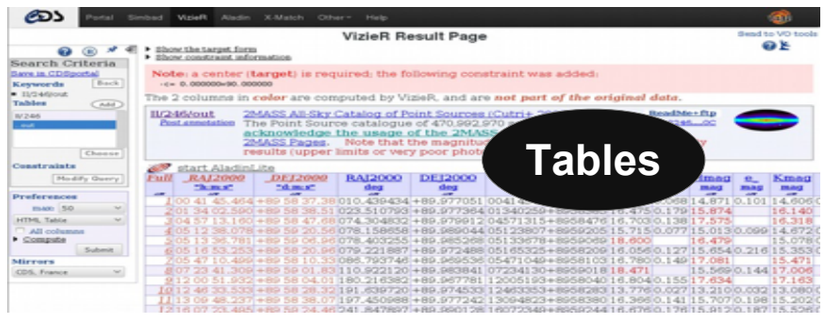
Planned for 2021:

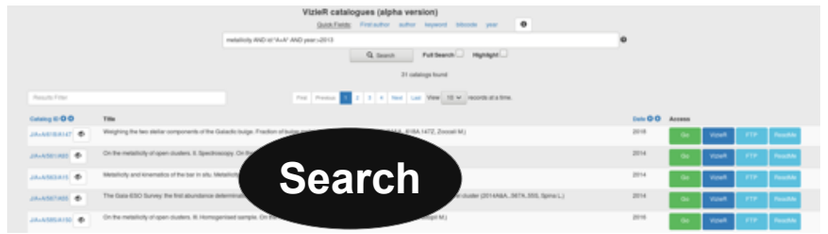
- ESO phase 3, new DRs
- GPS1+

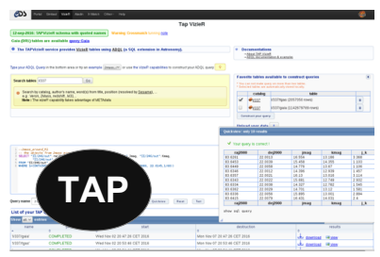
Accessing VizieR data

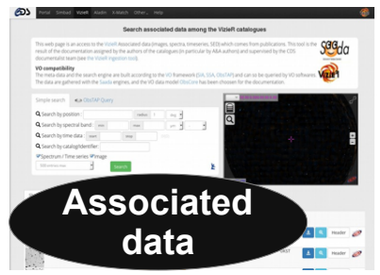
VizieR services and access modes: discovery, table access, visualization, photometry, associated data, astroquery, ...

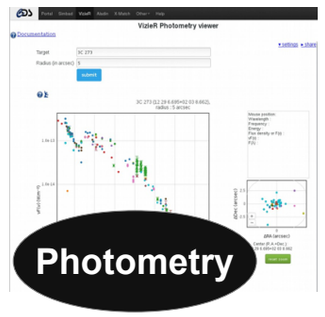


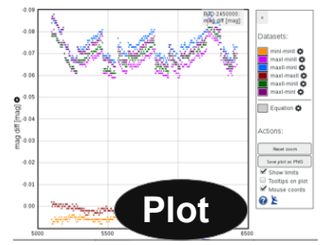




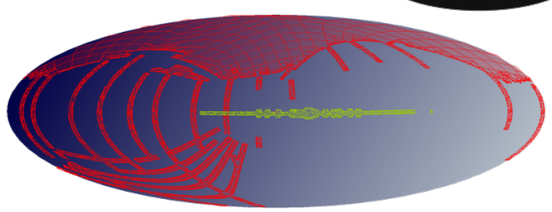


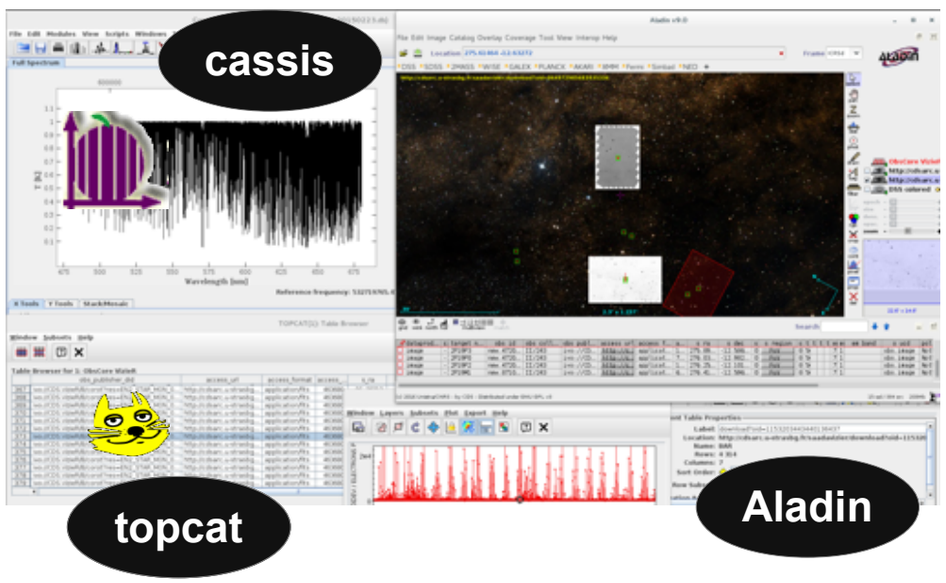






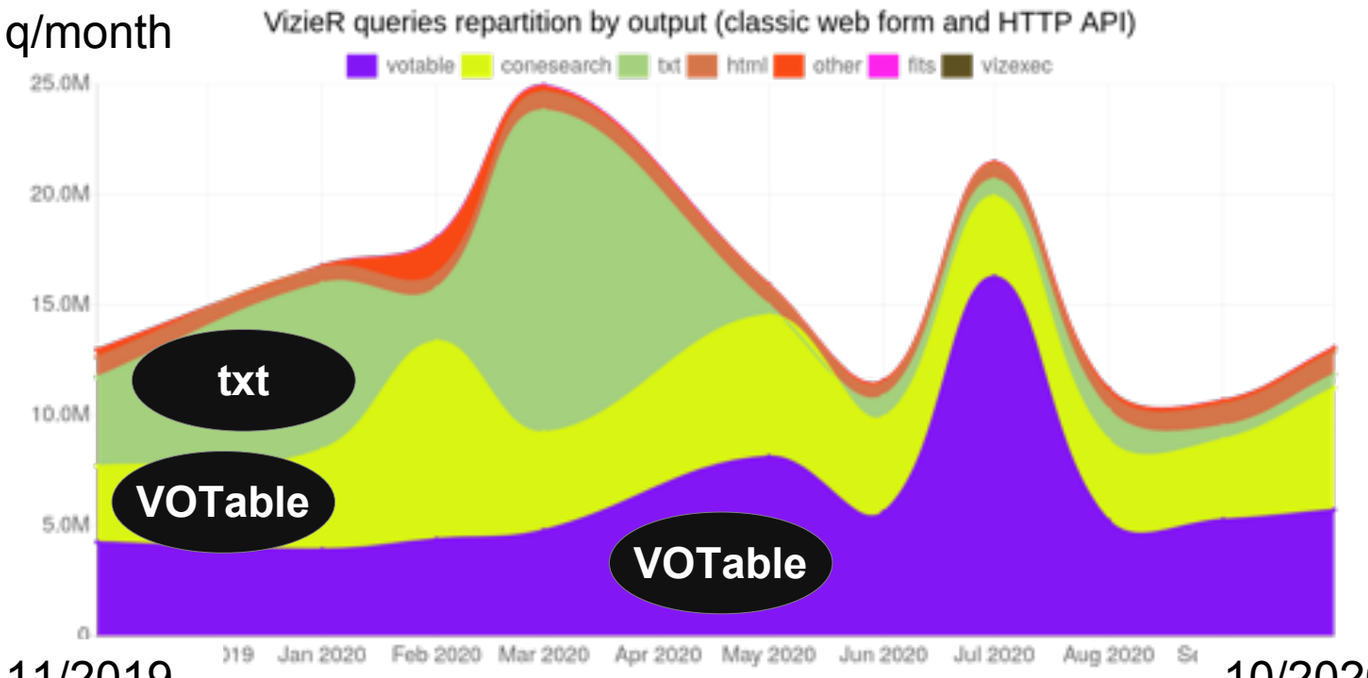
Derived products provided in CDS or data available through software, API..





VizieR 2020 usage statistics

VizieR Nov. 2019 – Oct. 2020 (from the CDS statistics collector)

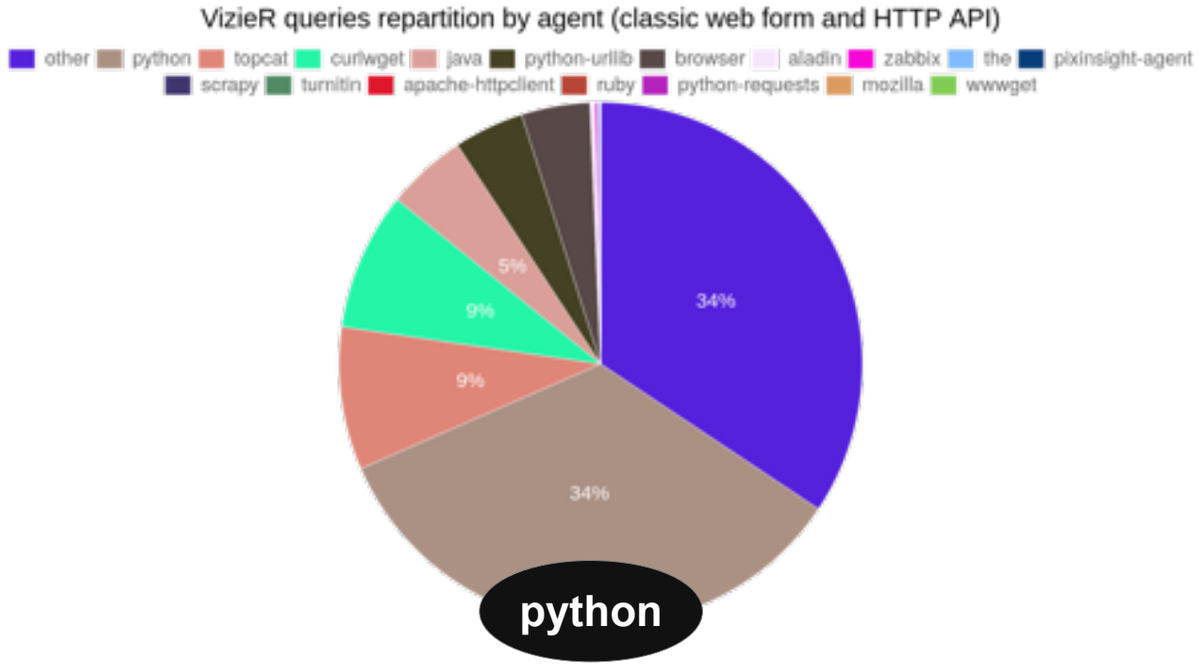


Total : ~520K queries/day

TAP : ~14K queries/day

Associated data (spectra/images): 845 queries/day

Dominated by Virtual Observatory formats (conesearch+VOTable)~70% of the queries



Dominated by Python and Topcat queries

Recent developments/deployments

- **Time domain**: since 2018, time columns are described with metadata (scale, frame, offset). Time column transformation to TCB/barycenter are now available in VizieR classic form and in photometry output.

- **DOI**s are now generated for AAS catalogues as well (only A&A in 2019). Vizier is the first public (i.e. non-private) creator of DOIs in number in France in 2020.

- The new **textual search engine** (ElasticSearch technology) is used by the VizieR services. Enables complex, flexible search, like ADS “Modern Form”. E.g. “first_author:Bai year:2019”

- Submission interface update:
  - Update of authors submission interface for a better integration in the VizieR workflow.
- Other actions in progress / development:
  - ReadMe autofill for most frequent data types
  - Renewal of the UCD builder
  - VizieR/CDS registry : OAIPMH service to index CDS services and VizieR catalogue int the VO registry

Outlook – 2021 and 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:
  - Homogenization of pipelines (large catalogs, journal catalogs)
  - Homogenization of global positional indexation
  - Dealing with large and larger datasets:
    - Exploration of large table access with PostgreSQL
    - Distributed database architecture (towards surveys with high-frequency temporal sampling, LSST-like)?
  - Extension of time domain management / use
    - Global time indexation => allow temporal cone search

Aladin:

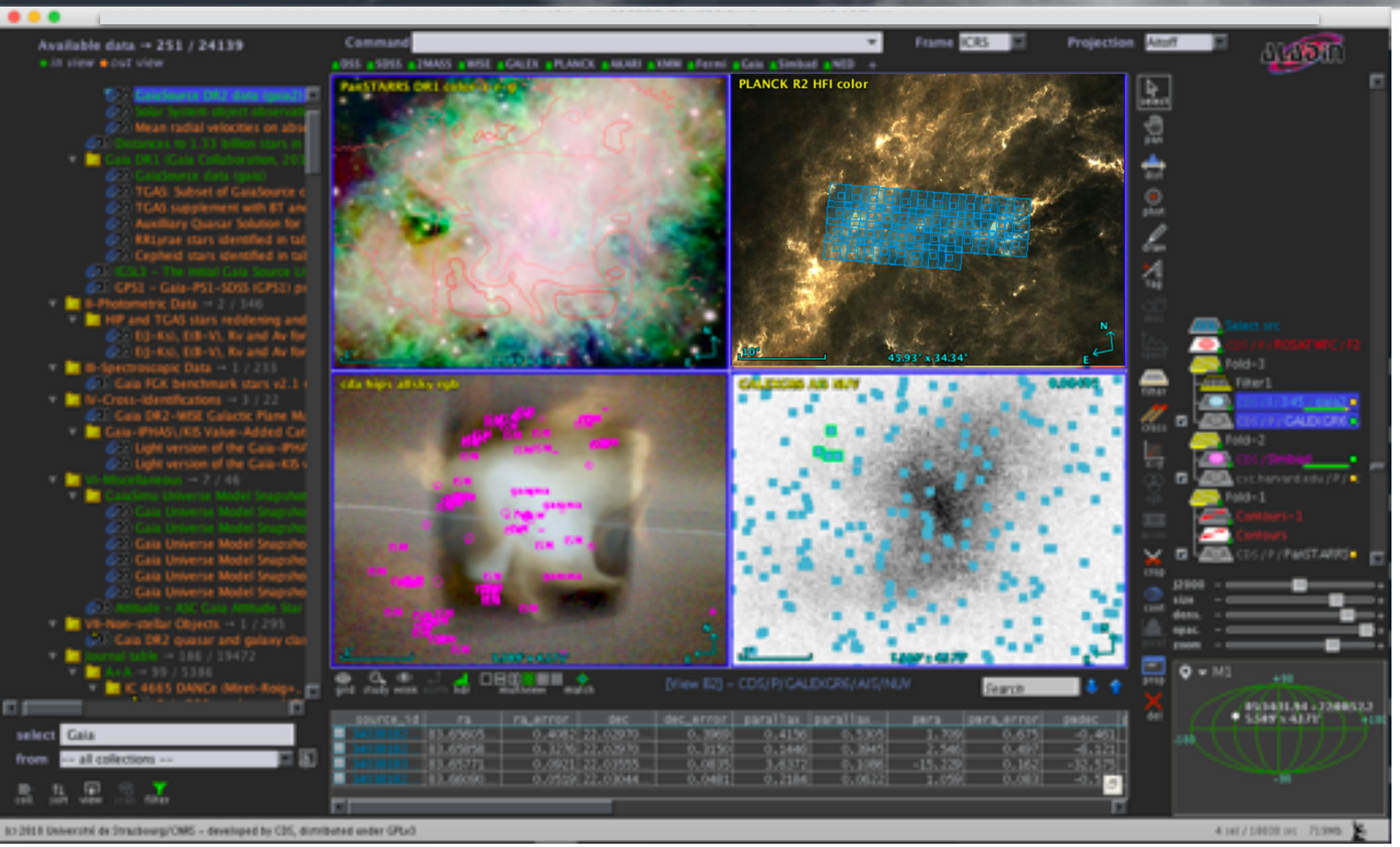
Highlights, statistics, perspectives

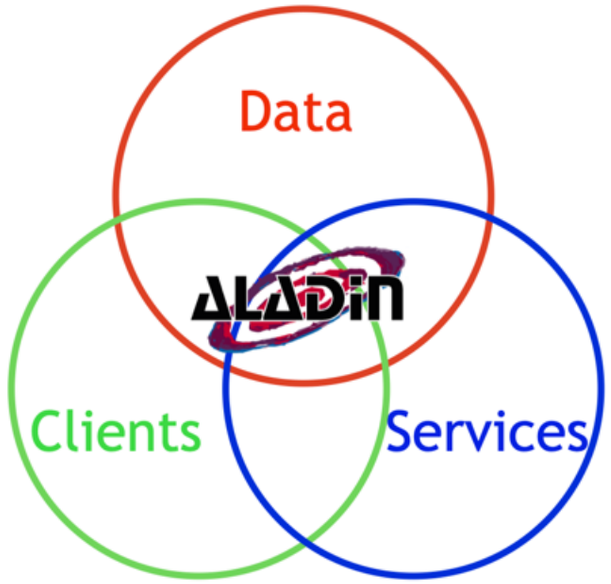
CDS council - November 30, 2020

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

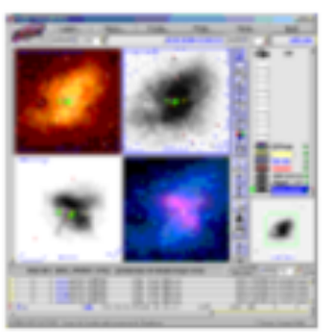
CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Aladin project





HiPS2FITS
HiPSgen, ...



Download
Aladin Desktop
on your machine

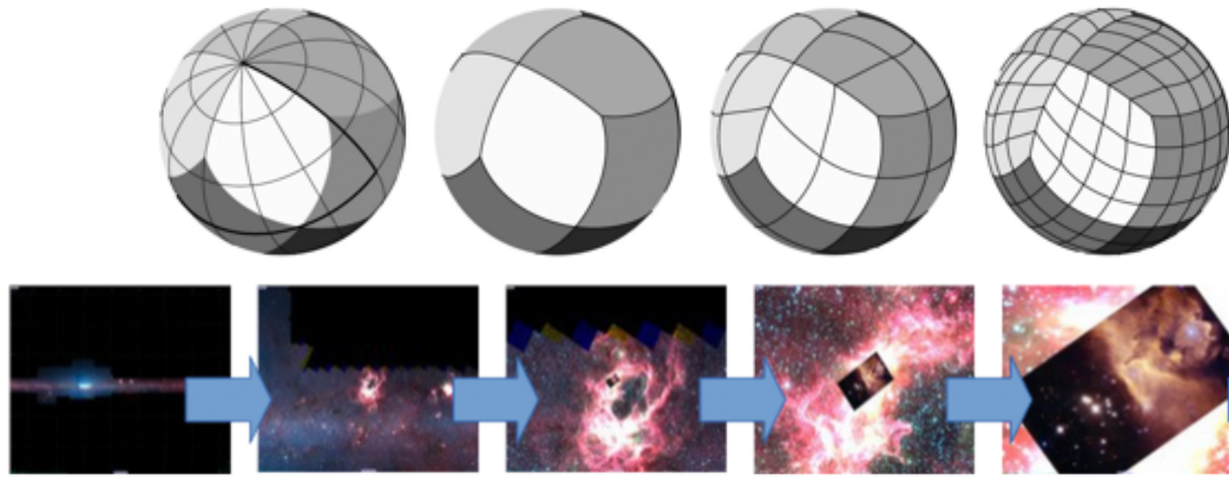
Preview with
Aladin Lite
in your browser

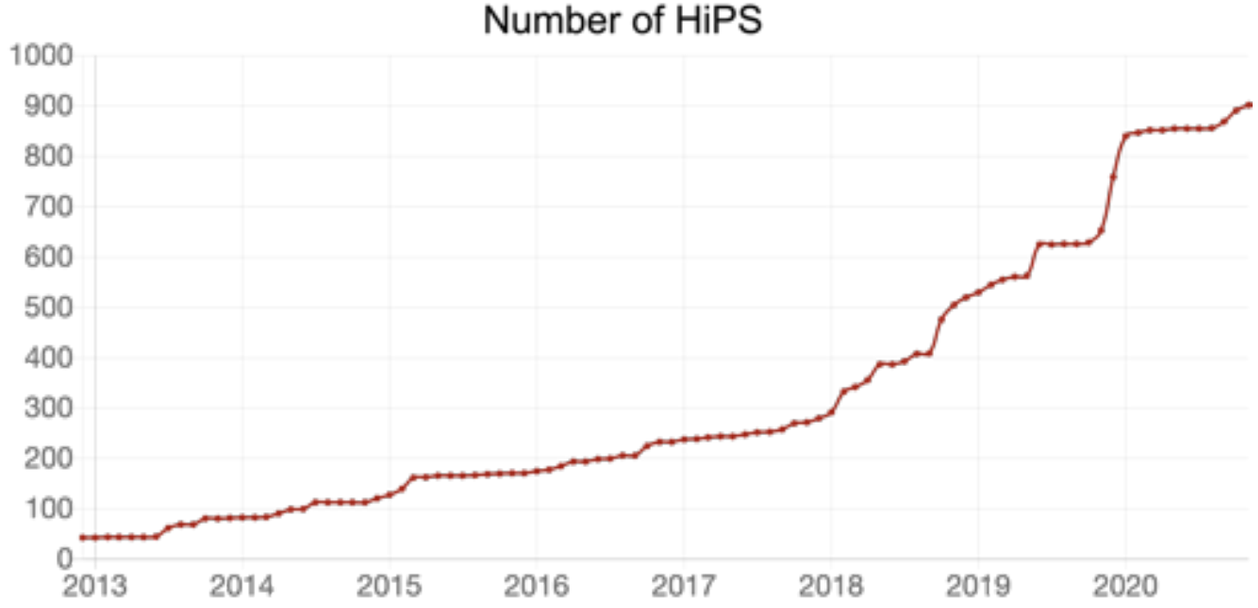
Aladin V11 release!

April 20 2020

HiPS ecosystem continues to grow

- **Hi**erarchical **P**rogressive **S**urveys
- 3 new HiPS nodes:
  ASTRON, JVO & CEFCA
- number of users up by 20%

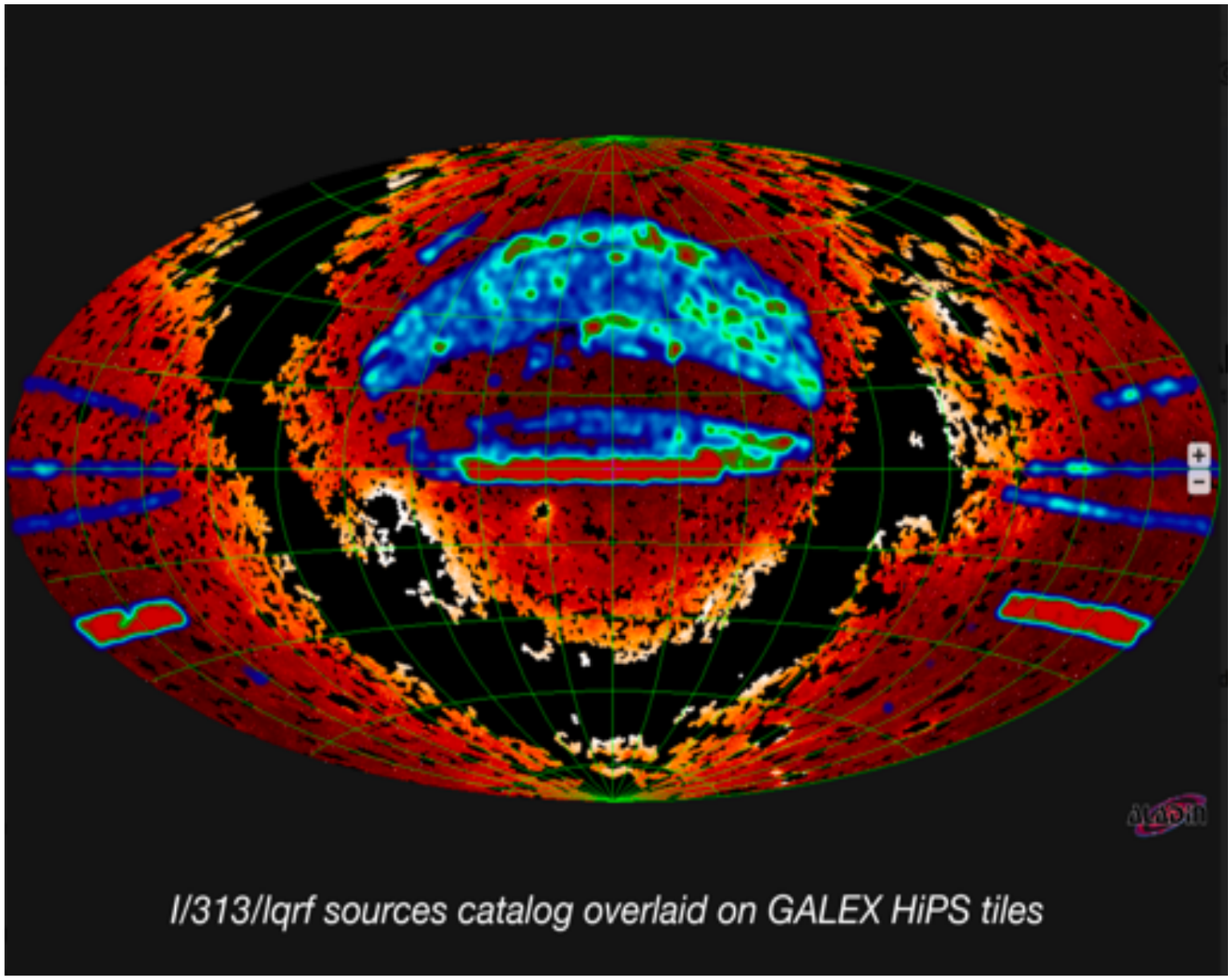




902 HiPS total

370 TB total

Aladin Lite v3 being developed



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

- WebGL & WebAssembly
- enable new projections (AITOFF, Mollweide,…)
- support the display of FITS tiles
- visualisation of catalogs up to 1M rows
- Colour blending of HiPS datasets
- same API for developers & integrators
- public release planned for 2021

Coronelli Globe

Mellinger HiPS

Aladin Lite

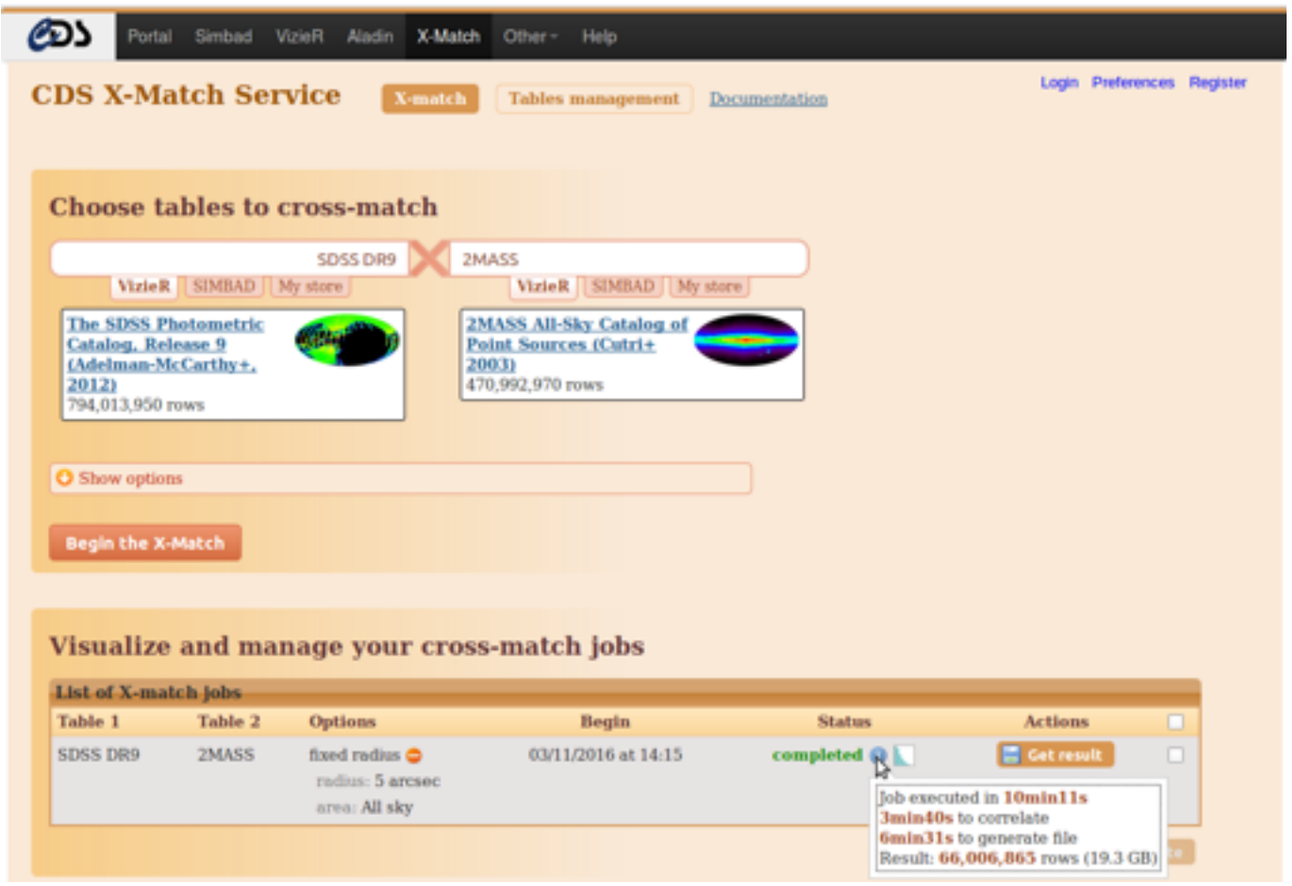
CDS Scientific Council
The X-match service

F.-X. Pineau and the CDS team

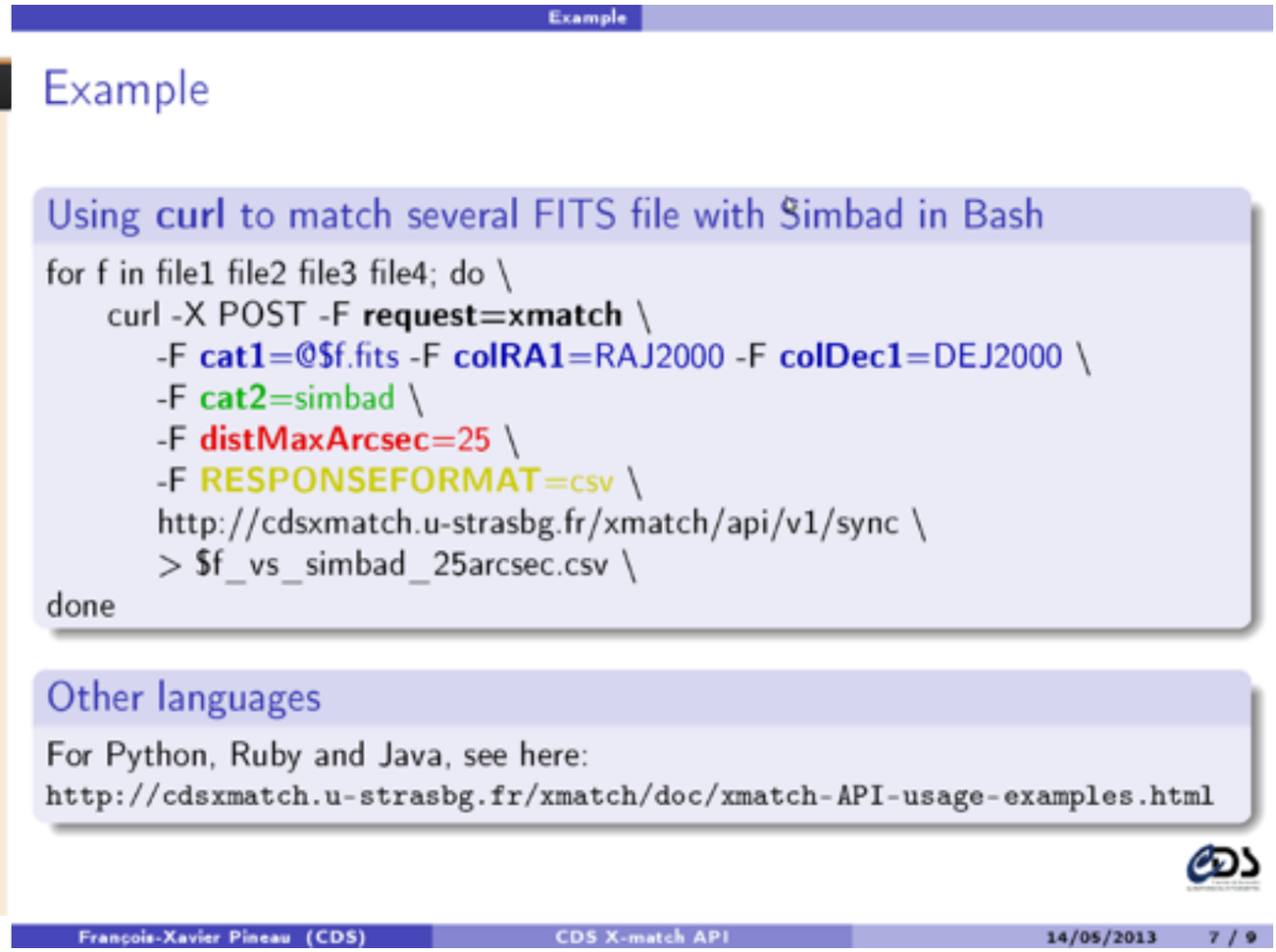
Virtual meeting, 30th November, 2020

CDS XMatch Service

- Very efficient cross-match of two (possibly large) tables
  - Any VizieR table and Simbad
  - User uploaded table
- Smooth operations in 2019-20 with current hardware and software

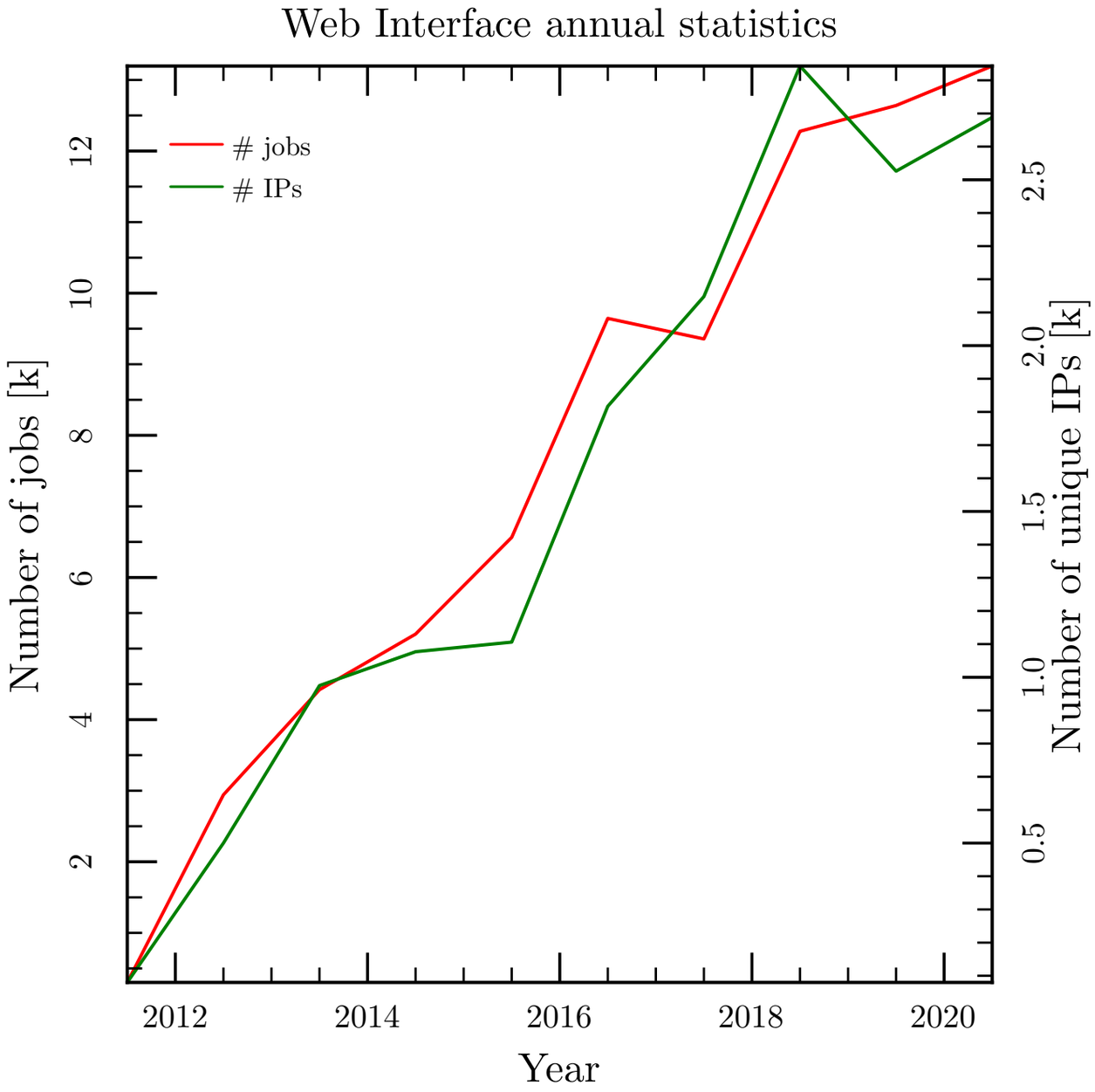


Web interface



Programmatic access (HTTP API)

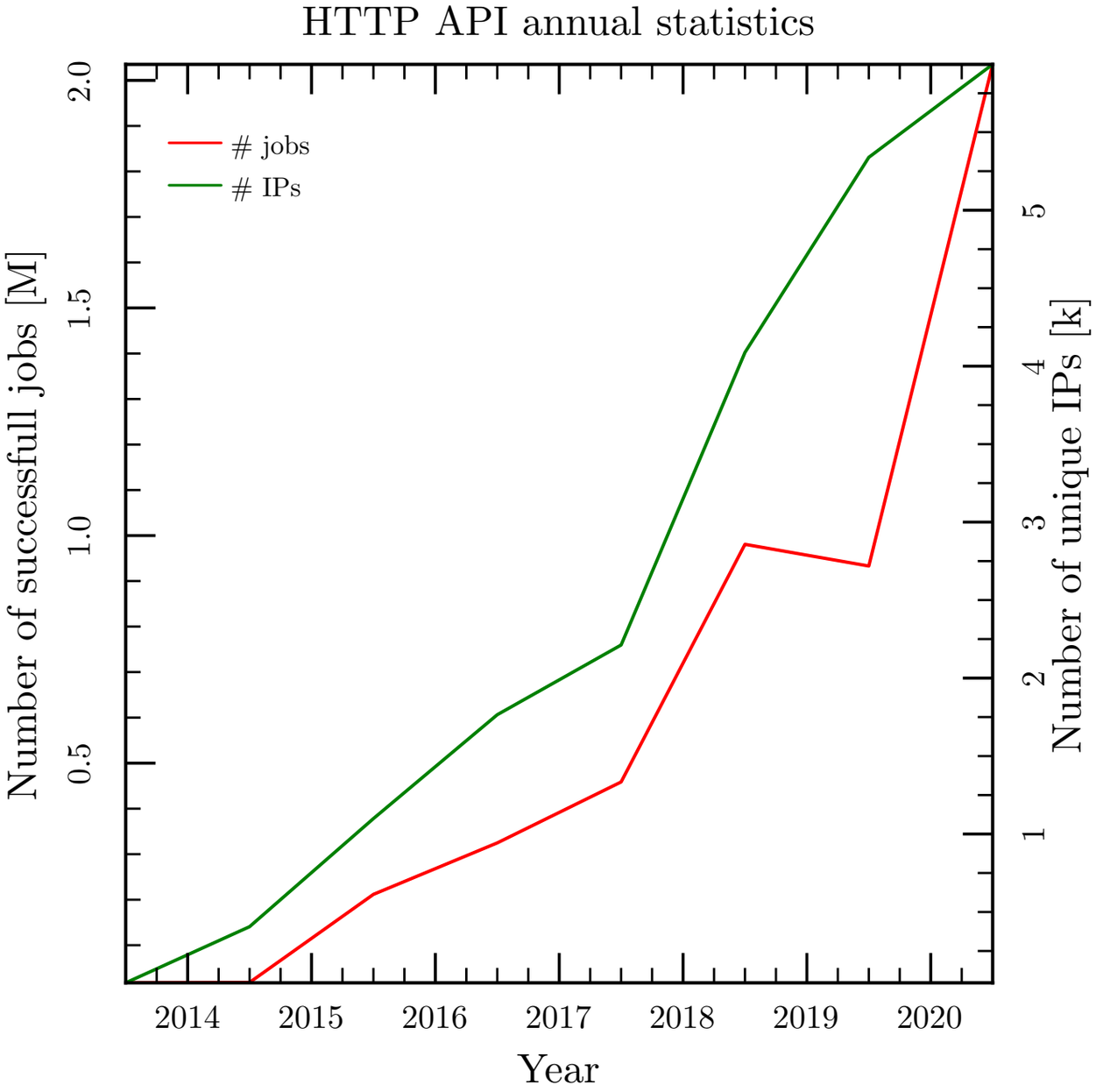
Annual usage statistics



Web interface

- #nJobs still slightly growing

Annual usage statistics



HTTP API (synchronous jobs)

- Usage still growing
- #nJobs: >**5000** jobs/day (+100%)
- #nIPs: > 700/month
- >**80** million associations / day

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)
- 2021:
  - Continue the development of the new cross-match prototype

R&D @ CDS

and various developments

André Schaaff on behalf of the CDS Team

CDS Scientific Council 2020

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.

- Technology watch to follow the evolutions is becoming hard in addition to the everyday work.

- 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 of the IT team through the presentation of the results.

Internship programme

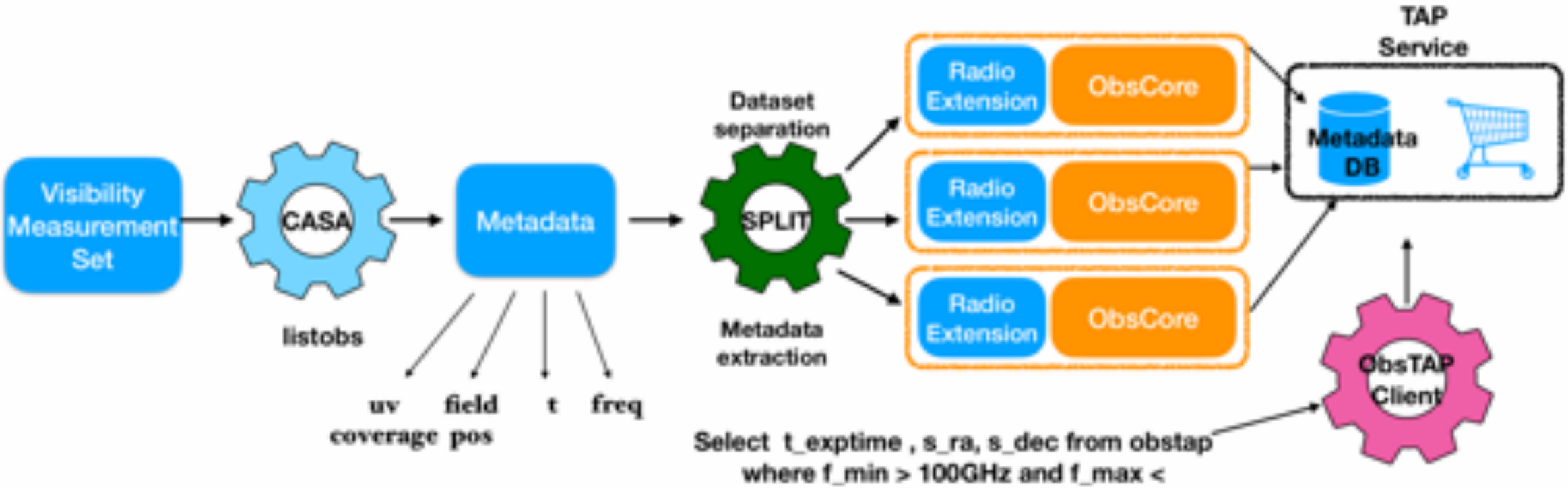
- 12 interns in 2020 worked with us on several topics, R&D and various developments > 3 years FTE

… in telecommuting this year !

- + 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.

IVOA (1) Radio Visibility extraction for discovery with ObsTAP



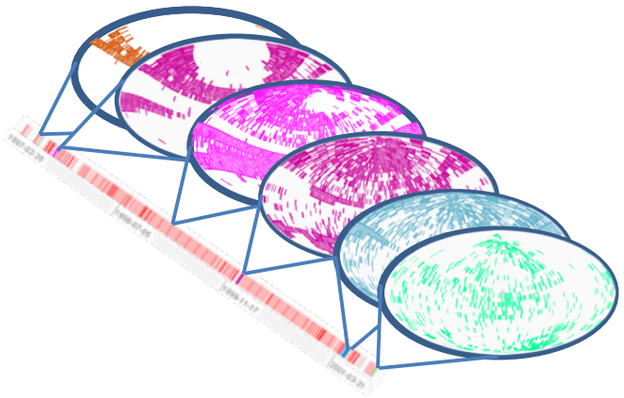
- ESCAPE H2020 / IVOA Radio IG
- Prototype to explore and organise radio visibilities in various datasets Using an extension of the IVOA ObsCore Specification.
- Goal is to allow ObsTAP query by criteria on spectral, temporal, spatial, polarimetric coverage.
- Split procedure under discussion with radio archives (LOFAR, EVN, Nançay, ATCA, etc.)
- The Radio Extension metadata is discussed in IVOA Radio IG.
- Poster at ADASS 2020

M. Louys, F. Bonnarel, K. Lutz, Y. Stein
Intern: Anaïs Egner (Université Le Havre)

IVOA (2), standards evolution, implementation, …

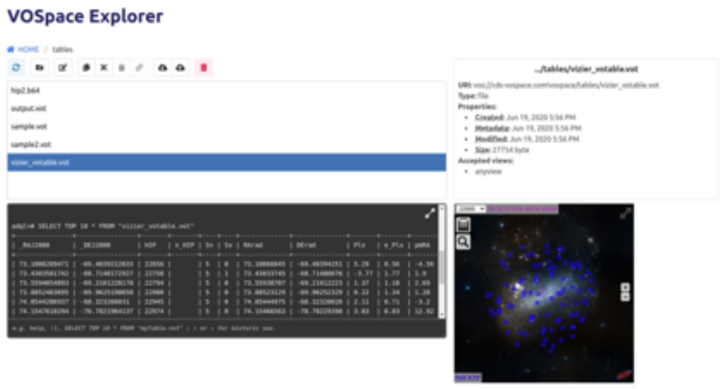
- Multi-Order Coverage Map was promoted to an 1.0 IVOA Standard in 2014 and was only based on Space. The introduction of time (driven by the IVOA Time Domain IG) was presented in the R&D talk during last year Scientific Council. Since then An effort was made to rewrite the IVOA document (2.0 Working Draft available since this Month) to introduce both Space MOC, Time MOC and Space-Time MOC consideration.

P. Fernique, A. Nebot, D. Durand (CADC) et al.



- VOSpace prototyping

G. Mantelet
Intern: Grégory Adam (IUT Schuman Illkirch)



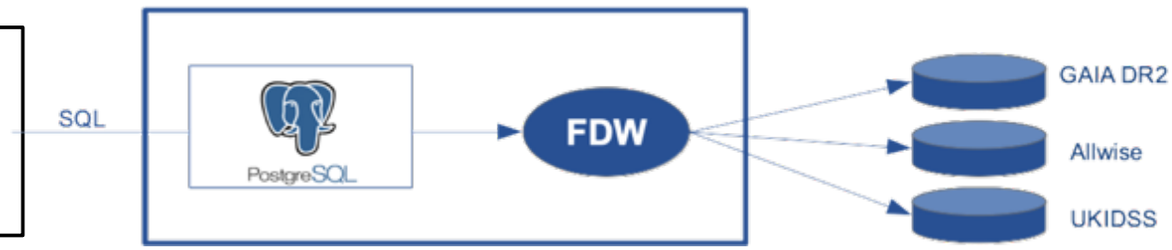
- Provenance of VizieR catalogues using the IVOA Provenance Data model (to specify the curation activity of a catalogue for VizieR "consumers".

G. Landais, F. Bonnarel, L. Michel (ObAS), M. Louys, M. Servillat (LUTH, Observatoire de Paris), M. Sanguillon (LUPM Montpellier)

VizieR Access to remote large tables

- 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
Intern: Alexandre Lepicier **(UTC)**
Adrien Vizier-Fontanesi **(ENSIIE Paris) in 2019**

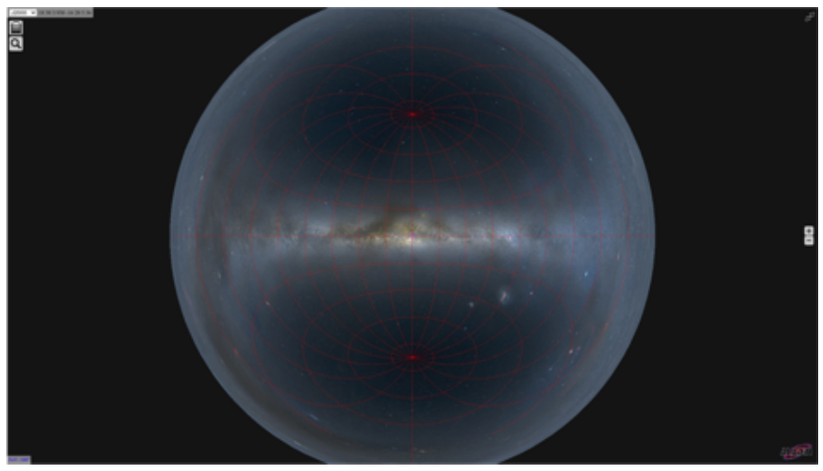


- + Development of a semantic analysis tool for the description (IVOA UCD1+ standard) of tables in astronomy.

S. Derriere, G. Landais
Intern: Louis Demange **(IUT Charlemagne Nancy)**

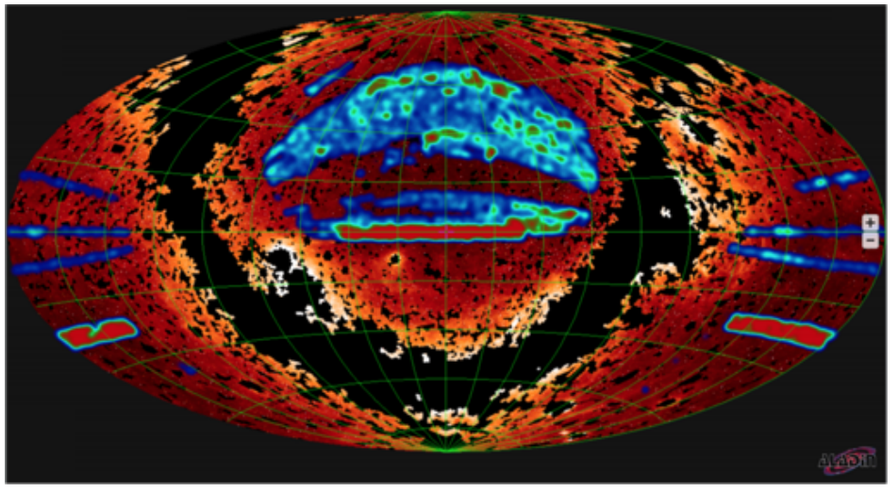
AladinLite 3

The new version benefits from WebGL which exploits the client hardware capabilities through OpenGL ES for a faster 3D rendering and provides new opportunities and performances close to a desktop application.

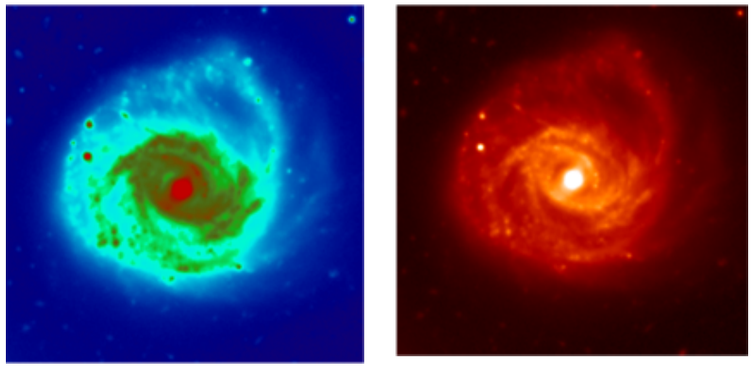


Mellinger HiPS Arc projection

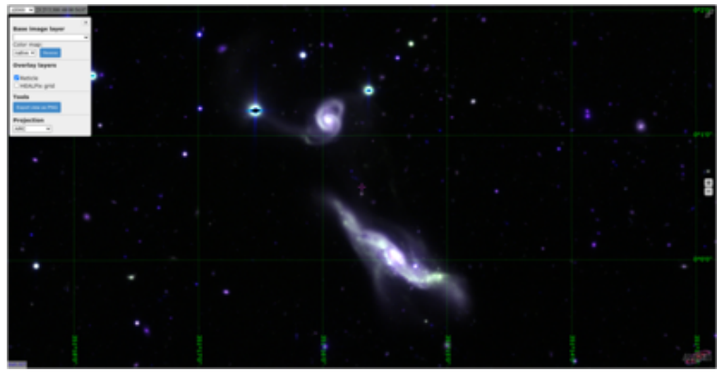
Oral at ADASS 2020.
WebCast available.



GALEX HiPS Survey aitoff projection with I/313/lqrf catalogue overlaid as heat density map



FITS values to colormaps

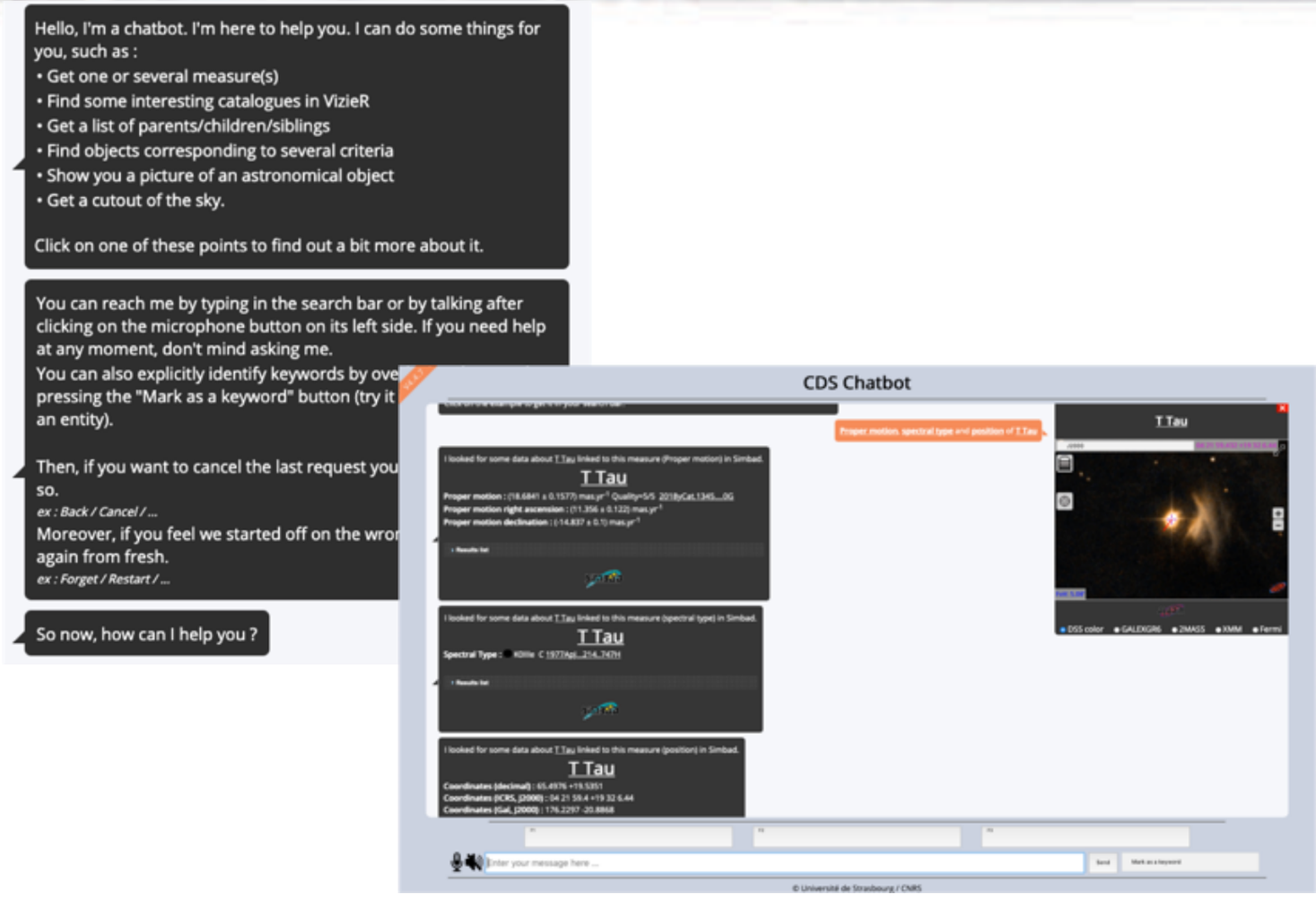


3 (red, green, blue) HSC HiPS composition

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 M 31 in optical

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

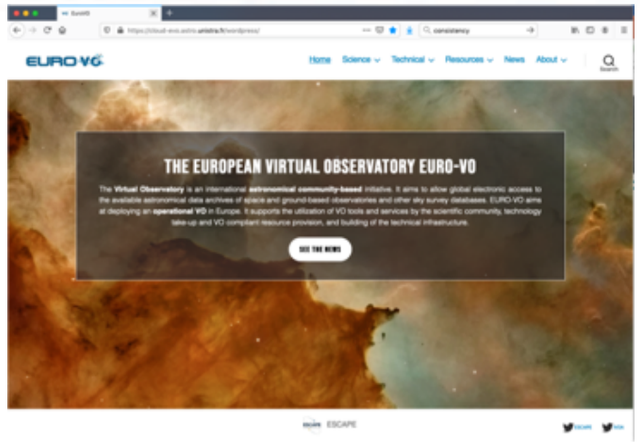
A. Schaaff, T. Boch, S. Derriere

Interns: Benjamin Monserand (UTBM)

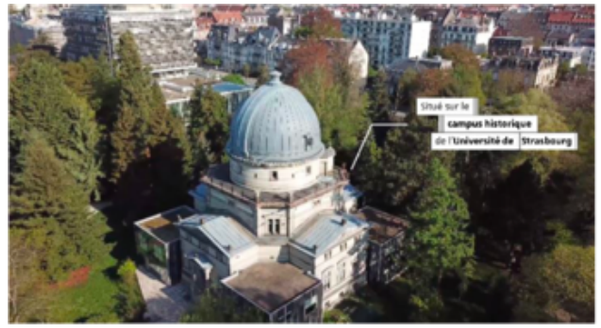
… and Pierre Sinnaeve (UTT), Antoine Herkens (IUT Belfort-Montbéliard), Alexis Guyot (IUT Dijon)

Quickly but not less important…

Reniewing of the euro-vo.org Website (hosted at CDS since 2013)



A. Schaaff, K. Lutz, M. Allen
Intern (communication): Catherine Menguy (IUT Saint-Dié-des-Vosges)



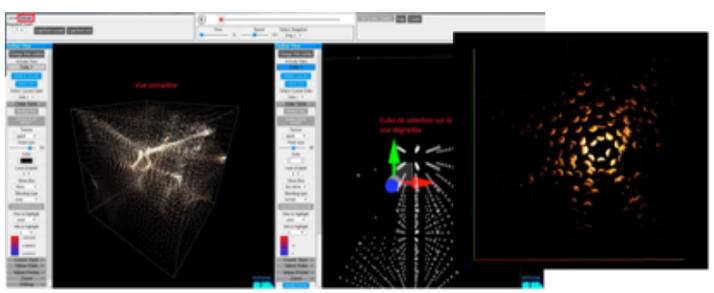
https://youtu.be/b4uZpxUgpZc

Video editing to create a presentation of the Observatory, a teaser and a valorisation of the Coronelli globe

https://youtu.be/lVNZmbTu4gg

P.-A. Duc, S. Derriere, A. Schaaff
Intern(communication): Odile Mallet (IUT Saint-Dié-des-Vosges)

Major evolution of JASMINE3D Visualisation (of large datasets) prototype to WebAssembly



A. Schaaff
Intern: Rova Rasoanaivo (IUT Saint-Dié-des-Vosges)

CDS Dashboard

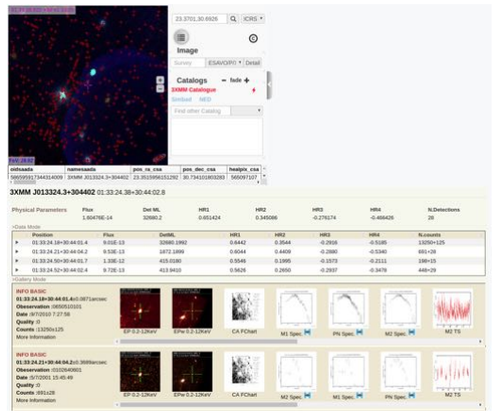
See Pierre's presentation

P. Fernique
Intern & short contract: Jean Miclo (IUT Schuman Illkirch)

Several collaboration topics (@ ObAS)

- With L. Michel (GALHECOS team), currently around Alix, a VO oriented user interface based on Aladin Lite

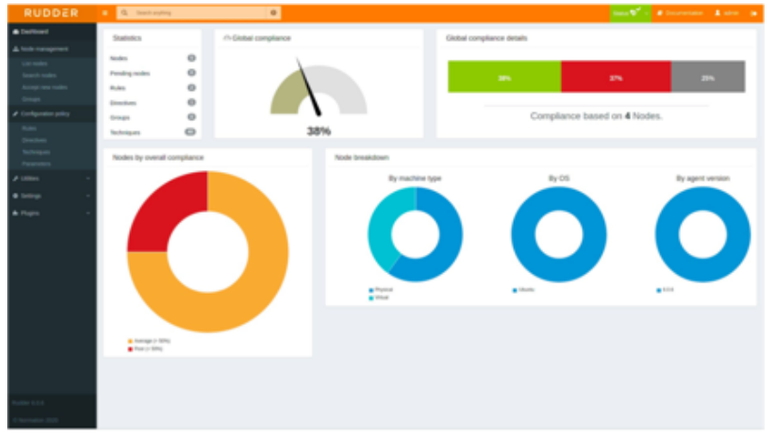
L. Michel, T. Boch
Intern: Serges ZOME (UTBM)



- With C. Saillard and T. Keller (Infrastructure team) to contribute to the ObAS IT infrastructure, this year an internship around the managing of the ObAS IT Infrastructure software consistency.

C. Saillard, T. Keller
Intern: Théo ERTZSCHEID (IUT Schuman Strasbourg)

Continuation as an apprenticeship



Future investigations

- As usual not Geek-driven, a continuous R&D effort to provide an added value to the data access & presentation mechanisms.

- Science platforms
  - hot topic (IVOA, ADASS, ESCAPE H2020 / EOSC, …)
  - A coordinated Science-driven CDS investigation should be on the rails in 2021.

- Efforts to make the data and services accessible to Everyone.

Conclusion

- A large 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.

Demonstration

Sébastien Derriere

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG