CDS Scientific Council meeting 2018

Summary of CDS activities 2017-2018

8 October 2018

1. Introduction

The 2017-2018 period has been a very active one for the CDS where we have pursued our core mission as a reference data centre for astronomy with productive, yet stressed, regular operations for the ingestion of reference data into our services. We have provided scientific support of the use of the services, and have made progress on the renewal of the internal architectures of the CDS services.

The publication of the most important, and largest, new astronomy data sets has been transformative for the CDS services in this period, with the Gaia mission second data release providing a global improvement of astrometric accuracy, and the PanSTARRS survey providing high resolution wide coverage reference images. A significant amount of effort was spent on this in 2017-18. These new data have been intensively used by the astronomy community.

CDS has continued to be highly visible in the astronomy community, in major conferences, workshops and training events. The documentation of the CDS services and scientific outreach materials have been improved and updated in this period, including our newly initiated presence on social media. Many collaborative activities have been pursued including visits to CDS partners, and participation in various scientific and technical projects. This has lead to an increased level of integration of CDS tools in other organisations, and has provided support to coordinated efforts toward global interoperability of astronomy data and services.

Planning the future infrastructure development and security of the CDS has been a very important topic in 2017-2018. A proposal to CNRS-INSU for a major renewal and upgrade of the main CDS data server has been successful. We are closely following the construction of the Data Centre at the Strasbourg University. The CDS disaster recovery plan has been reviewed and plans made for its formalisation, benefitting from the expertise of the new ObAS system engineer. Important upgrades of various CDS servers and other hardware have been made in this period, including the installation of a virtual server replacing several physical machines. The CDS R&D program has also pursued various topics related to operational actions and exploratory development work.

CDS continued its leadership roles in the IVOA, building and maintaining the Virtual Observatory framework for interoperability of astronomy data and services, with significant contributions in many of the Working- and Interest-Groups. CDS has also been active in the RDA and leads the RDA France National Node set up in March 2018. A new European project called ESCAPE, following the ASTERICS project has been successfully proposed in 2018 ensuring the future participation of CDS in data sharing developments for very large astronomy and astroparticle physics infrastructures in the new context of the European Open Science Cloud (EOSC), and incorporating both IVOA and RDA aspects. CDS has also applied for renewal of its certification by the CoreTrustSeal.

In this report we provide an update of the status of CDS in the national and European landscape in Section 2. Highlights of the 2017-2018 period are described in Section 3. The activities of the CDS services are described in Section 4, and Projects in Section 5. Section 6 provides the CDS responses to the 2017 recommendations of the CDS Scientific Council.

2. Status of the CDS in the National and European landscape

Important developments have occurred in 2017-2018 in the national and international landscape with respect to data sharing, and large infrastructures. Below we outline various topics that relate to the status of the CDS, and the effect they have on the CDS strategy.

French National Roadmap for Research Infrastructures

In 2017 we reported that the CDS had its status renewed as a **"Research Infrastructure"** on the National Research Infrastructure Roadmap established by the Ministry of National Education, Research and Innobvation (MESRI). This roadmap document was published in May 2018, with CDS included as an Astronomy and Astrophysics "Virtual Infrastructure", with strong European and International dimensions related to the International Virtual Observatory Alliance (IVOA) and our international partners. The roadmap recognises CDS' role as a driver of the VO at the international level, and coordinator of VO activities within Europe.

This roadmap highlights the importance of research infrastructures that produce, process and share data. All the infrastructures combined represent a total volume of ~540 Petabytes, which is predicted to increase by a factor 5 over the next 5 years. Indeed CDS has ~1 Petabyte of storage now and is expected to grow by this factor. The roadmap document recognises the necessary augmentation of the resources that will be needed to support the increased volume of data, and the digital infrastructure to support it (network, computing power, storage, archiving). The French strategy for dealing with data is described as being in line with the European Commission strategy for e-Infrastructures, namely the *European Cloud Initiative* and its two components the *European Data Infrastructure (EDI)* and the *European Open Science Cloud (EOSC)*. One of the objectives of this strategy is that data produced by research infrastructures should adhere to the ***FAIR*** principles, meaning that they are: ***F**indable, **A**ccessible, **I**nteroperable, and **R**eusable*. References for FAIR include the definitions from the FORCE11[1] (community of scholars, librarians, archivists, publishers and research funders), and also Wilkinson et al. 2016[2].

The CDS activities are inherently FAIR as the CDS reference data is: i) made *findable* via various interfaces and via the VO registry, ii) *accessible* via multiple interfaces, iii) *interoperable* via the use of standards, and iv) the CDS services make the data *re-usable* beyond the initial purpose of the publications or surveys they were derived from. Indeed CDS and astronomy have been providing FAIR data long before the emergence of the concept. The architecture of the VO is now also routinely mapped onto the FAIR principles. The adoption of this language to describe the CDS strategy helps the wider understanding of CDS activities and reinforces our role within the national roadmap and in the international landscape.

MESRI National Plan for Open Science

Another relevant reference is the MESRI National Plan for Open Science[3] published in July 2018. This again reinforces the Open Science concept of making research publications and data freely available, and it also emphasises FAIR principles. This plan announces the French support of the Research Data Alliance (RDA) for best practices concerning research data. The roadmap includes a point to "*actively contribute to structuring European data in the European Open Science Cloud…",* and other recommedations directly relevant to CDS such as the need to *"develop subject-based and discipline-specific data repositories"*, and the adoption of Open Data policy associated with journal articles.

---

[1] https://www.force11.org/group/fairgroup/fairprinciples

[2] https://www.nature.com/articles/sdata201618, Scientific Data volume 3, Article number: 160018 (2016)

[3] http://cache.media.enseignementsup-recherche.gouv.fr/file/Recherche/50/1/SO_A4_2018_EN_01_leger_982501.pdf

European Cloud Initiative and European Open Science Cloud (EOSC)

The European Cloud Initiative[4] is a high level European Commission policy to *"unlock the power of big data for open science",* and the European Open Science Cloud[5] is designed to "*bring together current and future data infrastructures"*. It intends to enable data sharing with open and seamless services to analyse and reuse research data to improve science. The descriptions of EOSC can be somewhat vague, but with significant levels of funding made available via European Commission Horizon 2020 programme in calls related to EOSC, this is obviously *the* major data initiative for research data in Europe which will define the landscape of research data infrastructures in the near to medium term. As such it is particularly important to CDS as a Research Infrastructure, and to our collaborations with large astronomy and astroparticle physics projects. It is important to note that EOSC is not only concerned with technical aspects but also strongly supports the development of human skills for *stewardship* of the data and services. This includes the recognition of career paths of the people who make it happen, and as such is of high relevance to improving the recognition of the work of the CDS staff.

The 16M€ ESCAPE project (described in more detail in Section 5) has been successfully proposed in such a H2020 call in 2018 (with our astronomy, astroparticle physics, and Euro-VO partners) in which CDS will lead a work package that will link the VO infrastructure with the EOSC. This is a development that will have an important effect on our future strategy as the CDS and VO services must be put into the context of the EOSC.

ESFRI Roadmap

The European Strategy forum on Research Infrastructures (ESFRI) published the update of its roadmap[6] in September 2018. This covers all fields of science and innovation, and the Physical Science and Engineering (PSE) category includes the astronomy infrastructures: SKA, ELT, and CTA. The document outlines the evolving role of research infrastructures, noting that they are increasingly becoming part of a connected system, and that EOSC will be an overarching project with a structuring impact on European science in particular to enable interdisciplinarity.

The Landscape Analysis presented in the ESFRI Roadmap (section 3 of that document) includes an analysis of "Big Data and e-Infrastructure needs". ESFRI adopts the FAIR data principles, reproducibility and Openness. The document notes that "EOSC will federate the most advanced data and service infrastructures, often directly built and supported by RIs". The long term investment that has been made to establish the astronomical Virtual Observatory is recognised in the following excerpt of the document:

"*In the PSE domain, astronomy has pioneered a global framework for FAIR data sharing which is operational and intensely used by the international community: ground and space-based observatories provide access to their data which can be reused for scientific aims different from the initial motivation of the research: a Virtual Observatory (VO) defines the relevant data standards as well as state-of-the-art data analysis tools. The VO shows the power of interoperability within a discipline to enable data and Commons to become an integrated Research Infrastructure. The international VO Allliance (IVOA) is expanding with the inclusion of astroparticle physics needs and*

[4] https://ec.europa.eu/digital-single-market/european-cloud-initiative

[5] https://ec.europa.eu/digital-single-market/en/european-open-science-cloud

[6] http://roadmap2018.esfri.eu

through the Astronomy and ESFRI Research Infrastructures cluster (ASTERICS) and planetary physics by[7] the Virtual Atomic and Molecular Data Centre (VAMDC)"

It is remarkable that the VO and ASTERICS come out at this level in the whole Physical Science and Engineering domain. And that the concepts become more broadly adopted, including another point made in the ESFRI analysis about the RDA:

"RDA is developing reference criteria and methods and a broader concept of virtual observatory/ laboratory enabling remote access to RIs spanning over the PSE domain".

This recognition of astronomy developments, strongly contributed to by CDS, has a reinforcing effect on the CDS strategy to continue to play a leading role in IVOA and building the interoperability framework alongside the large astronomy and astroparticle physics Research Infrastructures.

The different elements driving the evolution of the landscape of data sharing and large infrastructures have been outlined above. The significant developments over the 2017-2018 period appear to lead to a good level of coherence between the different bodies that define the landscape, in particular with the consistent adoption of the language of FAIR and Open data, and uniform support for the ambitious program being defined for the EOSC. This helps enormously for CDS to be able to describe its activities and strategies in these terms so that they are understood by a wider scientific community and the various technical and policy-maker stake-holders. The inclusion of CDS in a H2020 funded project to help build the EOSC, and the recognition of activities such as pursued by CDS in ASTERICS and RDA, puts CDS in a favourable position to continue to develop its future strategies in line with these large initiatives.

Certification of CDS

Another aspect of the status of CDS is the renewal of the certification following the Data Seal of Approval (DSA) certification that was initially done in 2014 for the 2014-2017 period, and the renewal of the CDS membership of the ICSU World Data System (ICSU-WDS). Here there have also been important developments whereby the two major certifying bodies have joined and have now launched a new certification organisation called the CoreTrustSeal[8] (CTS). Due to the timing of this transition to CTS, CDS has opted to renew its certification with the new CTS system (rather than a renewal within DSA in 2017). The CTS application for certification has been submitted in September 2018, and the details are provided in the VizieR service report (Section 4.2).

MENESR Complete Costs

In 2017 we reported that the MENESR led the 95 French national Research Infrastructures (TGIR & IR) in an exercise of calculating their complete costs. This exercise was again required in 2018. The CDS and observatory administration prepared all of the information, including calculation of the material/investment costs via the "ammortissement" option that considered all of our major hardware investments back to 2010. The reported annual cost of the CDS is 3.2 M€, made up of personnel costs of 2.2 M€, functioning costs of ~900 k€ (including indirect costs), and material costs of ~160 k€. This is higher than reported in the 2017 exercise, due to large IT purchases and the higher costs of contract personnel.

[7] We note an confusion of the role of VAMDC in this text, VAMDC is applicable generally to atomic and molecular data, not just in an planetary science context, and perhaps this text from ESFRI should have mentioned the EuroPlanet consortium

[8] https://www.coretrustseal.org

3. Highlights 2017-2018

Gaia DR2

On April 25, 2018, at just after midday, CDS published the ESA Gaia mission Data Release 2. This data release was eagerly awaited by the astronomy community, and the CDS release of the data provided unique capabilities that were heavily used by the community with some 6 million queries in the first month. CDS Gaia activities were done in coordination with Gaia DPAC. A special Gaia@CDS page[9] provides a direct link to the Gaia data in the CDS services (VizieR, Aladin and the CDS catalogue cross matching service) including an all-sky hierarchical view of the entire data set, and a CDS-generated flux map (shown in the figure below). CDS also provides visualisation of the 500,000 DR2 light curves, and the data have also been made available via Virtual Observatory protocols (IVOA Cone Search, and Table Access Protocol).

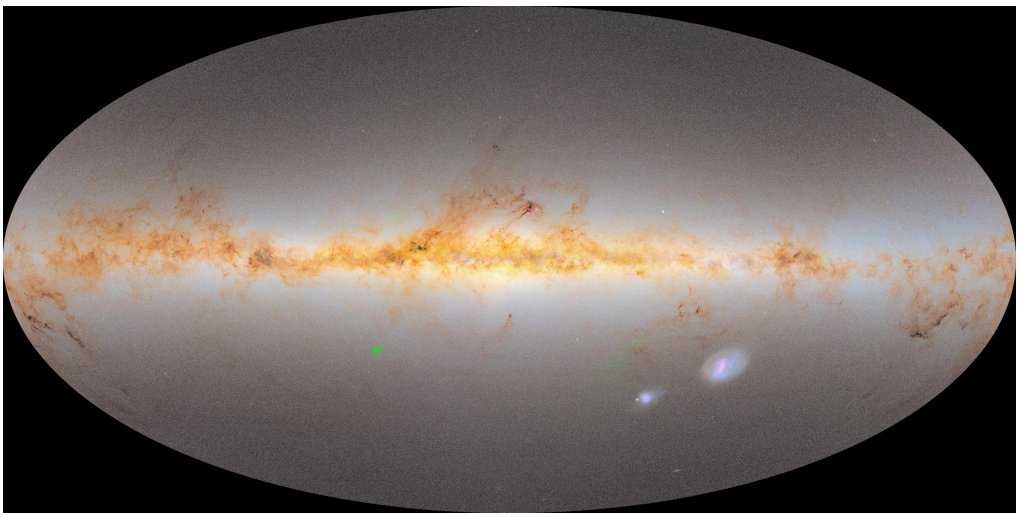

Fig 1. — Gaia DR2 flux colour map, built from G (green channel), Bp (blue) and Rp (red) fluxes in catalogue. Green areas correspond to stars without Bp and Rp measurements.

Gaia DR2 in SIMBAD

The cross-identification between the SIMBAD database and the ESA Gaia mission Data Release 2 has been successfully performed in June 2018, leading to a major improvement of the SIMBAD database with some 4.5 million objects now having a significant improvement in astrometric accuracy, and the addition of the Gaia measured parallaxes, proper motions and radial velocities. Coordinates of Gaia objects were calculated at epoch 2000.0 taking into account the proper motions in order to be compared to the positions in SIMBAD.

Aladin Desktop Version 10

A new major version of the Aladin desktop application, Version 10, was released in December 2017. This new version has a more modern “look and feel”, and many new features have been enabled via the implementation of IVOA standards. The Aladin Desktop interface is intended to be “full-featured” and allows for complex queries and supports our “power-users” as well as simple usage.

[9] http://cds.unistra.fr/gaia

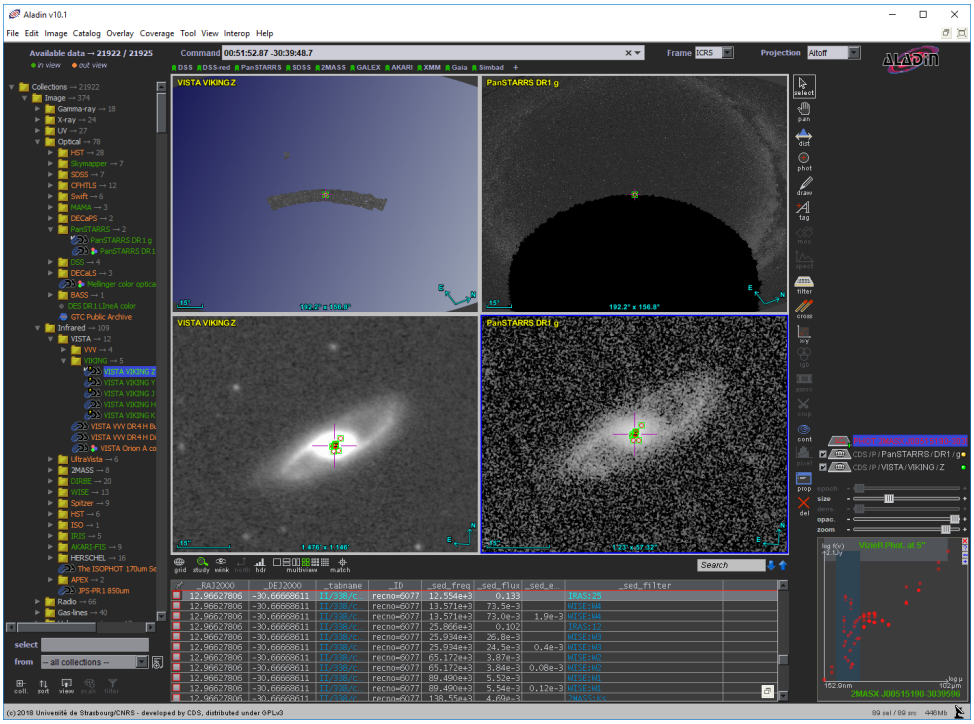

Fig 2.— Aladin v10 with HiPS VISTA-VIKING Z (left) compared to HiPS PanSTARRS g (right) with the photometry viewer result (VizieR, bottom right) on 2MASX J00515190-3039596 galaxy

CDS All-Sky Data project

CDS has made a successful proposal in the CNRS-INSU call for “mi-lourds” (medium sized) projects in 2018. The CDS “All-Sky Data” project enables a major update to the CDS data storage system, with 200 k€ obtained to purchase a 2 Petabyte storage system (1 Petabyte replicated for security). The new system will allow us to take into account the changing scale of astronomy survey data, in particular for our HiPS system. The system to be acquired in late 2018 is designed to be installed over two locations with one duplicate eventually in the Unistra Data Centre.

Documentation and Social Media highlight

A set of papers about the CDS and its data ingestion procedures have been written and published in the proceedings of the *Library and Information Services in Astronomy*[10] (LISA) conference hosted by CDS in 2017. These papers which highlight the work and expertise of the CDS documentalists and engineers are now available[11] with Open Access in the EPJ Web of Conferences published by EDP sciences, and form an important set of reference papers for CDS processes.

CDS now has a presence on social media with a facebook account[12], a twitter feed[13] and a YouTube channel[14]. These new lines of communication are helping us to reach new audiences.

---

[10] http://cds.unistra.fr/meetings/Lisa8/

[11] https://www.epj-conferences.org/articles/epjconf/abs/2018/21/contents/contents.html

[12] https://www.facebook.com/CDSportal/

[13] https://twitter.com/CdSportal

[14] https://www.youtube.com/channel/UCUESQl7rNupLIV_VcceE0Ng

European Project - ESCAPE

CDS is part of a successful proposal to the Horizon 2020 Work Programme INFRAEOSC-04-2018 Call – "Connecting ESFRI infrastructures through Cluster projects". The proposal was submitted in March 2018 following a rapid preparation, and was accepted in August, and will begin in early 2019 with a duration of 42 months. The project called **ESCAPE** (**E**uropean **S**cience **C**luster of **A**stronomy & **P**article physics **E**SFRI research infrastructures) leverages the successful experience of the H2020 ASTERICS cluster that brought together, for the first time, the astronomy, astrophysics and astroparticle physics facilities encompassed within the ESFRI roadmap. CDS leads one of the six work packages. The project has 32 partners and a total budget of ~16M€.

RDA - France

The RDA Europe 4.0 project started in March 2018. It is set up as National RDA nodes, with CDS in charge of the French node[15] in collaboration with Françis André from the CNRS Direction of Scientific and Technical Information (DIST).

Major data sets implemented in the CDS services

A number of high priority data sets (in addition to Gaia DR2) have been ingested into CDS services including PanSTARRS, SkyMAPPER and Dark Energy Camera surveys. A special effort has been made to include ESO public survey images into the HiPS system, with all bands of the VISTA VIKING and VVV surveys processed and available in FITS and jpeg HiPS formats.

Collaborations and Interactions with Partners and the Astronomy Community

The CDS has actively participated in many astronomy community events including the ADASS conference, IVOA interoperability meetings, project meetings of ASTERICS, AENEAS and Europlanet. An invited presentation and demonstration of the CDS portal was done at the UNOOSA/ASI Open Universe conference. A dedicated 2-day visit was made to ESO (Allen, Ocvirk, Fernique, Boch, Landais, Pineau) in November 2017, strengthening the CDS-ESO cooperation on large catalogues, and the use of CDS components in the new ESO Archive Portal. CDS also interacted with partners and scientific users as usual at the American Astronomical Society (AAS) Winter AAS meeting in National Harbour, Maryland in January 2018. CDS lead the installation of an IVOA booth at the International Astronomical Union General Assembly in Vienna, August 2018.

The CDS has also run a program of scientific training events in 2017-2018. We have made significant contributions to the following events:

- The ASTERICS DADI VO School[16], Madrid, 14-16 November, 2017
- The GAVO-CDS Gaia Data Access Workshop[17], Heidelberg 18-21 June, 2018
- A Doctoral school in Paris, 31 January, 2018

[15] https://www.rd-alliance.org/groups/rda-france

[16] https://www.asterics2020.eu/dokuwiki/doku.php?id=open:wp4:school3

[17] http://gaia.ari.uni-heidelberg.de/gaia-workshop-2018/

4. Activity Report for CDS Services 2017-2018

In 2017-2018 we have continued to use the ADS interface to track text citations of CDS services. The ADS interface allows counting of the number of papers in which the CDS services are cited in the text of the paper. **In the calendar year 2017, 665 refereed papers cited the word SIMBAD, 418 the word VizieR, and 74 the word Aladin** in reference to the respective CDS services. This represents a marginal increase for SIMBAD and VizieR compared to 2016.

Brief reports on the CDS services are given below, and more details will be provided in presentations by the service teams. A presentation on the scientific activities of CDS research staff will be made in the meeting.

4.1 SIMBAD

The content of the SIMBAD database continues to grow, the table below indicates the status on September 21, 2018. Some 12,000 new references (journal articles) have been ingested, and the processing of the main journals has been relatively smooth, despite technical problems with MNRAS. The growth of the number of articles processed by SIMBAD is shown in Fig. 3. We note that most objects are ingested into SIMBAD via tables from journal articles which have first to be ingested in VizieR. The number of tables of objects to be ingested in SIMBAD has increased by about 50% between 2017 and 2018.

SIMBAD Content

	2013	2014	2015	2016	2017	**2018**
Objects	~7,342,000	7,556,225	7,998,221	8,493,230	9,298,005	**9,639,763**
Identifiers	~18,162,000	18,563,653	22,322,732	23,553,608	26,799,877	**32,649,863**
Bibliographic references	~285,000	294,449	308,588	323,689	336,179	**348,341**
Citations of objects in papers	~10,000,000	10,749,766	12,126,329	14,352,859	16,169,095	**17,616,348**

In terms of managing the flow of incoming information, the initial processing of journal articles which identifies objects in the text and small tables of articles, is completely up to date (although stressed). The processing of objects in larger tables, many of which have to be first ingested in VizieR, currently has delays up to 2yrs (where a ~1yr timescale is preferable). The generation of acronyms is currently at a rate of about 15 acronyms per week, keeping up with the incoming flux, and catching up on a backlog of about 670 references (with a realistic prospect of decreasing this backlog by about 150 references per year).

The usage of the SIMBAD service remains very high with an average of 502,000 queries per day in 2018 as shown in the table below and in Fig. 4. An updated global distribution map is shown in Fig. 5. The query rate appears to be relatively stable over the past 5 years. The peak query rates have been observed to be higher in 2018, getting up to 2 million queries per day, but some of these peaks are due to inefficient bulk queries.

SIMBAD Service Use

	2013	2014	2015	2016	2017	**2018**
Queries/day	520,000	506,000	500,000	510,000	552,000	**502,000**

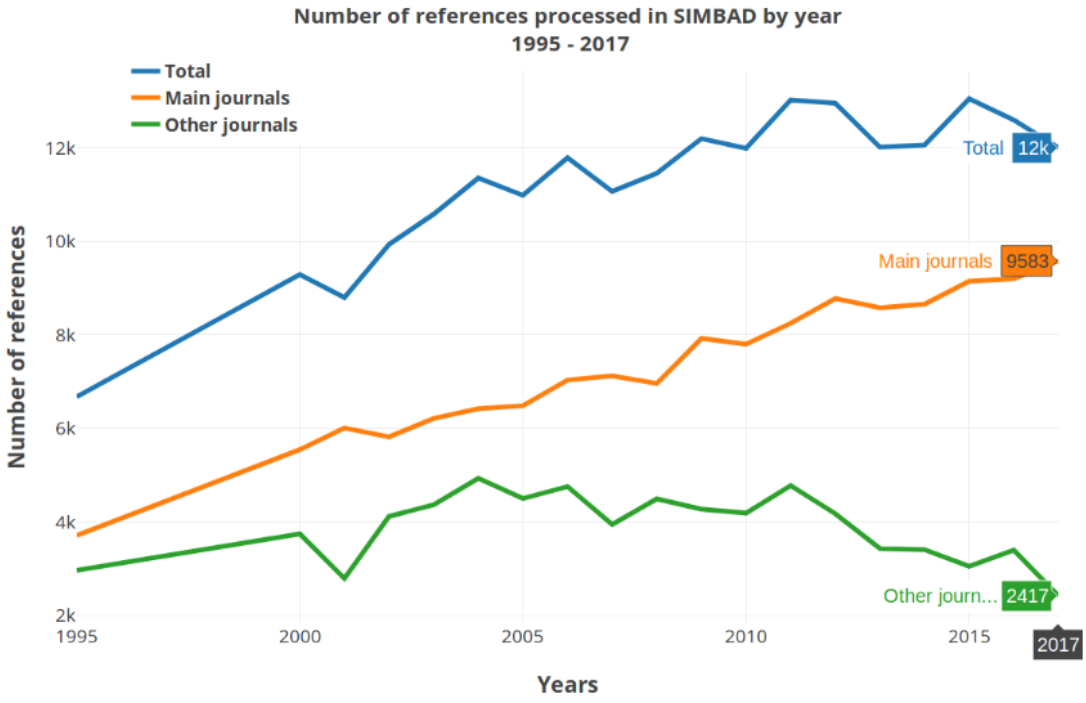

Fig 3.— References processed by SIMBAD 1995-2017

The major event for the improvement of scientific content of SIMBAD in 2017-18 was its cross-identification with Gaia DR2 in June 2018, only two months after the public data release. The criteria to perform this major operation were carefully studied and optimised. Essentially, the astrometry had to be sub-arsecond already in SIMBAD to avoid random matches, and the agreement between SIMBAD and Gaia DR2 positions had to be better than 1". Objects with neighbours within 3" in SIMBAD or in Gaia DR2 were discarded.

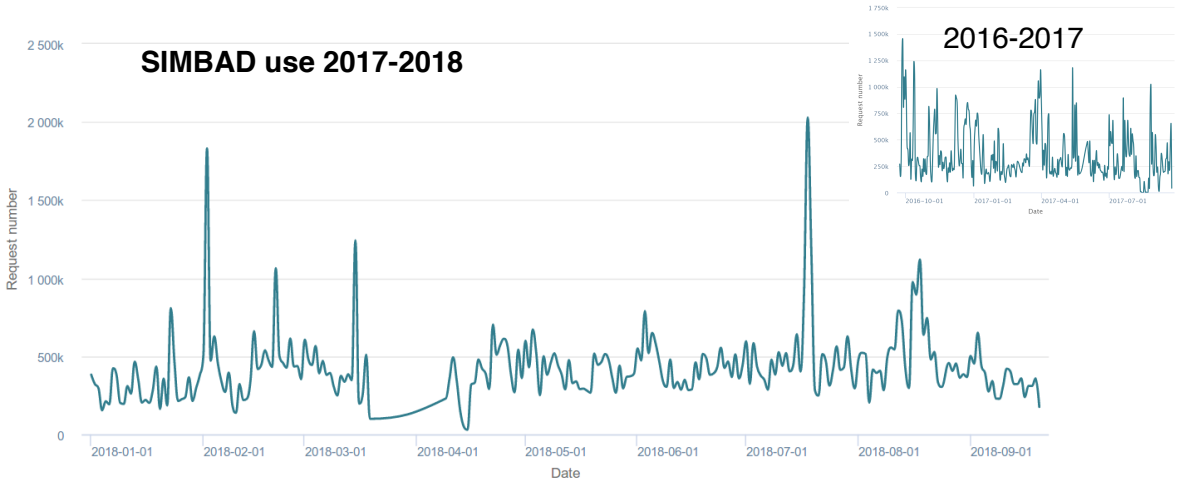

Fig 4.— SIMBAD queries in 2018

Out of the 6.5 million objects in SIMBAD that met the initial criteria on the astrometry, 4.5 million of these have been successfully cross-identified with Gaia DR2 : 4 million stars (including 85 % of high proper motion stars), and half a million galaxies (including AGNs). Most of the SIMBAD objects that were not recovered were either too faint to be detected by Gaia, or they were located in crowded regions. We consider this result to be a great success, and the fact that such a high rate of cross-identification has been achieved must be recognised as the result of more than 10

years of efforts to improve the coordinates and cross-identifications in SIMBAD. SIMBAD is now significantly improved by Gaia DR2, to the benefit of the CDS users, and also to the internal CDS processes which are now more efficient due to better coordinates from Gaia DR2. (Further details can be found in the CDS news item[18] about the cross-identifcation of SIMBAD and Gaia DR2.)

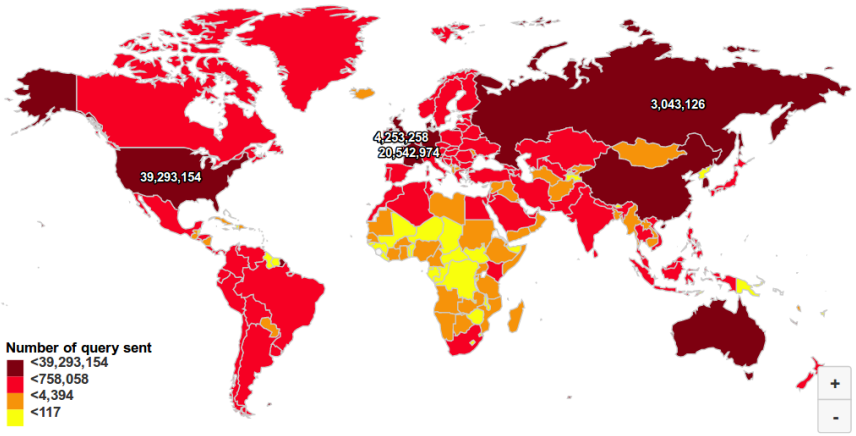

Fig 5.—Global distribution of SIMBAD queries in 2018

Other general evolutions in Astronomy that are having a big impact on SIMBAD include the improvement of spectral capabilities of observatories and missions, and the fact that samples of "rare" objects are now becoming large. Improved spectroscopy means that many more objects receive a "priority 1" label when they are detected by the SIMBAD ingestion procedure, a distinction that was more varied 10 years ago (in part leading to the 50% increase in the tables). In terms of "rare" objects, it has to be noted that for some classes, the numbers are now getting huge — for instance, there are now eight times more spectroscopically identified white dwarfs than there were 15 years ago; 39,000 compared to 5,000. The total number of confirmed AGN in SIMBAD is now about 402,000, but it will very soon be increased by almost 200,000 as we are ingesting the SDSS-DR12 catalogue of QSOs (2017). Similarly, there are now 105,000 eclipsing binaries and cataclysmic variables — this number will also soon be multiplied by a factor of 4 as we are ingesting the latest catalogue of OGLE towards the galaxy bulge.

Following the deep review of the SIMBAD architecture undertaken in 2017 (and reported last year) we have started important updates to the internal connections between the different components of SIMBAD (database, servers and ingestion tools). The structural changes are designed to enable a more simplified set of interfaces (APIs) so that developments and improvements can be done in a more modular way. This work in underway with a contract engineer who started in May 2018.

As reported last year, work on a major update of the DJIN tool was started in late 2016. DJIN is used by the documentalists to process the astronomy journals with the main task of identifying astronomical objects (names, acronyms and identifiers) in the text and tables of articles, an essential step at the beginning of the chain of CDS processes. DJIN operates on PDF files, and uses different parts of SIMBAD including the Dictionary of Nomenclature. This work has proved to be much more difficult than expected. A contractor who was hired to work on this project has now finished his contract. A revision of the project to define the next steps is now underway.

[18] http://cdsweb.u-strasbg.fr/news.php?fn_mode=fullnews&fn_incl=0&fn_id=702

4.2 VizieR

The standard VizieR treatment of catalogues has been in smooth operations throughout 2017-2018. The VizieR Content table below indicates the growing number of catalogues in the database.

VizieR Content

	2013	2014	2015	2016	2017	**2018**
Number of Catalogues in the VizieR database	11,579	12,691	14,065	15,485	16,528	**17,673**

In terms of use of the service there was an average of 368,000 queries/day in 2018. This is an increase compared to the 326,000 queries/day of 2017, partly pushed up by the publication of the Gaia Data Release 2.

There has been a change in the contract personnel working in the VizieR team. A documentalist, S. Guehenneux, left CDS at the end of February 2018. The loss of manpower is fortunately mitigated by the contribution of a new contractor T. Pouvreau, hired in March 2017, who successfully completed her training as a VizieR documentalist by the end of 2017.

VizieR Usage

	2013	2014	2015*	2016	2017	**2018**
Web Service Queries/day	600,000	530,000	300,000	380,000	326,000	**368,000**
Associated Data Service Queries/day					270	**80**
VizieR TAP service Queries/day					5,250	**3,700**

Gaia DR2 release has been a key event of the last 12 months. As a partner data centre in the Gaia Data Processing and Analysis Consortium (DPAC), CDS received the DR2 data one month in advance of the release date, and ingested them and made them public on April 25 at the same time as the official ESA archive and other partner data centres. It proved to be a much anticipated dataset: 2 million requests have been performed on Gaia DR2 catalogue in VizieR in the first week after the release (see Fig 6.). This figures rises up to 10 million requests in the first three months. TOPCAT and astroquery (Python scripts) have been identified as the most prominent clients.

Other large catalogues ingested during the reporting period include the PanSTARRS DR1 catalogue, Gaia distances (Bailer-Jones et al. 2018), GPS1 (Gaia-PS1-SDSS) proper motion catalog and two ESO public survey catalogues: VMC DR4 and VST ATLAS DR3.

Managing the load of queries on the VizieR service

The CDS services are operated on the basis of free and anonymous access. While this context is the most favourable for guaranteeing the impact and reproducibility of scientific results, some drawbacks have become apparent in 2017-18: there have been cases of queries that overload the VizieR system, rendering the main CDS VizieR site unresponsive for a number of hours.

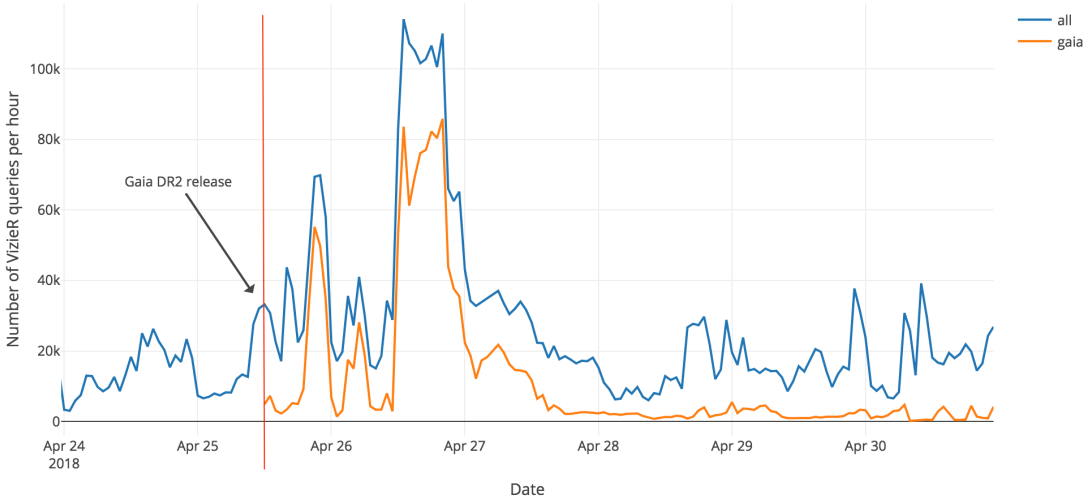

Fig 6.— VizieR queries per hour around the Gaia DR2 release date

The most recent incident involved a user performing several hundred thousand queries in a short amount of time, querying several large catalogues simultaneously within a large radius. This last incident, taking place in late September 2018, resulted in perturbations of the service for the users but also for CDS staff, and it required blacklisting of the IP address responsible for the queries. Identifying the problem and fixing it took a significant amount of the time of CDS engineers, and blacklisting an IP is not in the long term a satisfactory solution. CDS staff has been treating these events on a case-by-case basis, as a long term, more systematic solution has not been found.

Clearly, the free anonymous paradigm within which VizieR operates has pros and cons. Strategies for mitigating abusers have been considered but they also have their drawbacks. For instance, one approach could be to let VizieR remain anonymous for simple, light queries, but have the user login or authenticate for queries involving more than, say, 1000 rows of upload file, and/or a radius of more than a few dozens of arc seconds. While this sounds reasonable, it has a profound impact on the ecosystem and services using CDS data, since all data-heavy software downstream of CDS would need to adapt to account for this authentication process. An alternative, used by ADS for instance, is to use an API key, allowing only a fixed number of queries per day, while any further query will return nothing. This may not be an adequate mechanism for VizieR, given its role in operating observatories and data analysis alike. Finally, modifying our services in one of these directions will require a large amount of work from our developers. The alternative of course is to keep running as we do now, and accept that surges and downtimes may happen, in cases where we have not been able to see the problem arise quickly enough, and that these events require the urgent intervention of CDS developers. This may be the price to pay to have CDS services remain fully anonymous and unlimited in access. Altogether this is a complex issue with possibly far-reaching consequences and we would like to use the visit of the council to ask for advice.

Another aspect is that we may need to enable a prioritisation of access to guarantee that queries originating from CDS staff working on database curation will be able to perform nominally and complete, even during high server loads.

VizieR Associated Data Service

In 2016 VizieR opened a major new feature, the VizieR Associated Data Service for non-tabular data. This service includes an ingestion pipeline and the metadata mappings that are necessary to take into account the description of spectral, time-series and image data. Authors may provide

these data, and the ingestion process (supervised by documentalists and engineers) maps the description to the standard IVOA ObsCore data model. These data are discoverable through a web interface, and also through IVOA protocols.

This service is operational and we have been ingesting “associated data” that comes with newly arriving catalogues, but with a priority on ingesting the existing associated data already in the 17000+ catalogues already in VizieR. This is a huge undertaking and we started by processing the catalogues with the largest volumes of associated data and working our way downwards in size. This cautious approach allowed us to build experience with the process and to understand the resources needed to operate it. The service contains currently about 6.650 million fits files, mostly LAMOST spectra (319,174 non-LAMOST data products). At the moment however, the added workload of the ingestion procedure is still holding back the systematic ingestion of newly arriving associated data. The new service received about 80 queries/day, which is disappointing and a sharp decrease since last year. Clearly, the lack of advertisement is hurting the usage of the service. Communicating to the community about the service takes time and manpower too, and although communication was meant to be a focus point for the past year, we have not been able to do as much as we wanted. This service has come a long way, is operational, but there is still much to do and it is not clear that systematic ingestion can be performed with the current resources.

It is worth noting that the provision of non-tabular data associated with publications in VizieR, fully available in the Virtual Observatory, is aligned with the recommendations of the National Open Science Roadmap on the adoption of an Open Science data policy for publications, and on the provision of FAIR data. This contribution to Open Science is providing data from the so-called “Long Tail”, with a quality guaranteed by its association with a refereed paper, in the framework which also hosts the data from the large telescopes, with the same FAIRness capacities. As pointed out by C. Borgman in her book “*Big data, little data, no data*”, in general data from this Long Tail is simply not available although “little data” can be just as valuable as “big data”.

Digital Object Identifiers

The VizieR ingestion pipeline now routinely includes the DOI of published articles. In addition, a landing page was created for catalogs with DOIs to comply with the expected implementation of DOIs. Further agreements are being put in place with INIST, in order to mint DOIs for the catalogues themselves. Indeed, the catalog is a product distinct from the article, but needs to have its own DOI, referring of course to the one of the original article to which VizieR dataset was attached.

Time Domain

In VizieR, data homogenisation has always been both a challenge and an engine for progress. After coordinates systems and photometric systems, CDS staff has worked to equip VizieR for describing time systems in a consistent and systematic way. Several (6) new metadata tables have been created in order to describe catalogue time systems in terms of scale, reference frame, and representation, following Rots et al. 2015 (A&A 574,36). The determination/generation of these metadata by the VizieR documentalists is becoming part of the routine VizieR ingestion pipeline. This much-needed upgrade will support the development of time-domain astronomy, and a number of interesting new functionalities are planned to use this new feature.

Radio - Specfind

We are undertaking a special effort to update the CDS *Specfind* tool for the curation of radio spectra in VizieR. This is being supported by a CDS postdoc, Y. Stein. The new revision is expanding the number of radio catalogues from which data are extracted to contribute to the spectra, from 100 to ~200 catalogues, and this is expected to double the number of objects for which we can provide radio spectra.

Certification - CoreTrustSeal

In 2017-18 we have prepared the renewal of the certification of the CDS VizieR service. There has been a major development in the status of the relevant certification bodies, namely the ICSU World Data System (WDS), and the Data Seal of Approval (DSA[19]), whereby a new certification organisation has been launched, called the CoreTrustSeal[20]. The DSA has been merged into the CoreTrustSeal in 2018.

The CDS VizieR service has been certified by the DSA since 2014 for the period (2014-2017) and this status came up for renewal in 2017. With the change in organisation it was decided that the CDS would apply directly to the CoreTrustSeal in 2018.

A significant level of preparation was necessary to update the CDS documentation in order to describe the VizieR service in terms of the CoreTrustSeal application criteria. The VizieR team has made a major update of the "VizieR Processes" document (publicly available on the VizieR organisation pages[21]) which is used as the main reference for the CoreTrustSeal application. "VizieR Processes" describes the processes and pipelines of VizieR and how these are managed by the CDS staff. This description is done using the terms of the CCSDS[22] Open Archival Information System (OAIS) organisational model.

The CoreTrustSeal application involves a top level description of the CDS as a data repository, plus responses to the 16 requirements for the certification. Each requirement involves 1-2 pages of text, references and a self-assessment. The 16 requirements are:

1) Mission/Scope, 2) Licenses, 3) Continuity of Access, 4) Confidentiality/Ethics, 5) Organisational Infrastructure, 6) Expert Guidance, 7) Data integrity and authenticity, 8) Appraisal, 9) Documented storage procedures, 10) Preservation plan, 11) Data quality, 12) Workflows, 13) Data discovery and identification, 14) Data reuse, 15) Technical infrastructure, 16) Security

The self-assessment is a scale from 0-4, to indicate *"0 - not applicable"* to *"4 - fully implemented in the repository"*. The CDS self-assessment indicates 4 for most items. A self assessment of *"3 - in the implementation phase"* was indicated for the *Licenses* requirement, reflecting the need for some improvement. A self assessment of 3 was also provided for the *Data discovery and identification* requirement, indicating that the major improvement of creating DOIs for VizieR catalogues is being implemented. The CoreTrustSeal application was submitted in September 2018, and we await the evaluation.

Internal management

An interface for the internal follow-up of VizieR data ingestion (for use only by CDS staff) similar to a ticket system has been developed, and has been in operation for more than a year now. The interface works well and is being upgraded when required by the team to improve or specify functionalities.

Mirrors

In 2018 the Indian VizieR mirror (IUCAA) underwent a major upgrade, so that it now contains a full copy of all catalogues, in particular all the very large catalogues.

---

[19] https://www.datasealofapproval.org/en/

[20] https://www.coretrustseal.org

[21] http://cds.u-strasbg.fr/vizier-org/

[22] http://www.ccds.org

4.3 Aladin

The CDS Aladin service involves the Aladin Desktop application, the Aladin Lite web version, the underlying collection of images and data cubes in the HiPS system, the data ingestion processes, and a vigorous development program of new features and techniques.

HiPS, the Hierarchical Progressive Survey system developed by CDS and now an international standard, continues to be an important structuring development with new scientific functionalities in Aladin Desktop and Alain Lite largely being enabled by the conceptual and technical approach defined by HiPS. Furthermore the adoption of HiPS in the astronomical community is growing fast, with heavy use of the HiPS network (in which CDS provides the main node) via CDS and externally developed client applications.

In terms of usage we note another explosion in 2018 in the number of HiPS queries per day. The table below[23] shows the use by CDS Aladin and Aladin Lite. An important change is the large increase in the use of HiPS by other clients (Stellarium[24], FireFly, K-stars) and XMM-SSC which brings the total usage up from 307,000 to 770,000 queries/day with an audience of 170,000 hosts/ month. There is also a large volume of queries coming from the process of HiPS data being mirrored to other HiPS partner sites.

Aladin usage: Aladin & Aladin Lite

	2012	2013	2014	2015	2016	2017	**2018**
Queries/ day	13,208	56,053	160,103	266,047	276,000	283,000	**307,000**
Audience: hosts/ month	8100	8441	39,083	54,656	61,582	79,602	**66,941**

In terms of the data content in the Aladin service, a significant effort was made in 2017-2018 to collect, generate and ingest HiPS versions of a high number of important image surveys, including the very largest recent *'Tera-pixel'* surveys which are in high demand from the CDS users (PanSTARRS, SkyMAPPER, DECAM, ESO, +). This brings the total number of HiPS data sets to 575, with a data volume of 203 TB (see table below).

Aladin content: HiPS image data sets

	2012	2013	2014	2015	2016	2017	**2018**
HiPS data sets	81	128	175	236	325	380	**575**
HiPS Volume	19 TB	30 TB	45 TB	50 TB	105 TB	137 TB	**203 TB**

Appendix A shows a table of the largest data sets that have been included in 2017-2018. This reflects the scientific and technical choices that have been made. The PanSTARRS DR1 g-band has been made available as FITS and JPEG HiPS data sets, with a colour composition made from the g- and z-bands. This has proved to be an excellent reference data set for faint objects, and

[23] We have revised the way we report the "audience" with a figure now provided in terms of hosts/month, which is what is actually measured, rather than a conversion to a daily rate.

[24] https://stellarium-web.org

further bands will be computed with the timing of future data releases to be taken into account. All bands of the ESO VIKING and VVV public surveys, and all bands of the SkyMAPPER DR1 have been fully ingested. The creation of the HiPS for the ESO and SkyMAPPER data have been helped by a CDS visit to ESO, and the visit of SkyMAPPER to CDS. Selected bands of the DECam surveys (DECaLS DR5 and DECaPS) have been ingested. A number of large HiPS data sets have been created by external partners, notably SWIFT by HEASARC, BASS by China-VO, Dark Energy Survey by NOAO, and HERSCHEL by ESA.

By taking on the challenge of ingesting the largest data sets we have purposely pushed our system to its limit, allowing us to learn where the main difficulties lie in handing such large data sets. We have gained precious experience from this - we have overcome various technical issues involved in the internal transfers of large data volumes, and the temporary storage of data sets across multiple physical systems during the generation and ingestion phases. The current storage system is now close to capacity with ~100 TB allocated for surveys to be ingested in late 2018 and early 2019. Fortunately we now have secured the funding for the first phase of the expansion of the main storage system with the CDS All-Sky Data project, so we expect to continue this approach of ingesting the most important reference surveys.

Other important new HiPS data sets include the Gaia flux-map generated by CDS (shown in the highlights) as part of the publication of Gaia DR2 in the CDS services. Also we have published a special CDS data product derived from the SIMBAD bibliographical database, the sky density maps of astronomical objects cited in publications as a function of time (1850-2018) and by object type.

As shown in the 'highlights', a new major version of the **Aladin Desktop** application, Version 10, was released in November 2017 (following a pre-release at the ADASS meeting in October 2017). The new version has a modern 'look and feel' and has a large number of new features that have been made possible by the implementation of IVOA standards: TAP, MOC, Datalink. "Data Discovery" is a core concept in the new version, with the development of the interactive "data tree" that dynamically shows the user what data (images, 3-d data cubes, catalogues) are available in a given region of the sky, with the possibility of filtering and sorting the list in many ways. This tree provides access to data at CDS, but also external data that is made available via the VO registry. The user feedback obtained from the community via various training events, VO schools, workshops, and demonstrations at conferences (AAS, ADASS, IAU) has been very positive. Usage statistics show that 75% of Aladin users have moved to Version 10, and the number of "actions" performed via Aladin Desktop has increased by 25%. A relatively small number of bug reports have been attended to.

In terms of documentation, a full manual[25] has been produced in French, and a much needed English translation is awaiting the person-power to do it. Various CDS tutorial instructions have been updated to Aladin V10. Five short youtube videos have been produced highlighting the "Aladin data collection tree" (introduction and advanced), "Aladin SIMBAD pointer tutorial", and two tutorials on the 'Using the Aladin stack".

Since the release there have been further developments to support the IVOA standards relevant to the discovery and access of multi-dimensional data cubes. The standards are the *Simple Image Access Protocol version 2.0*, and the *Server Side Operations for Data Access 1.0 (SODA).* One of the innovations here is the development of an image cut-out facility based on HiPS, that is currently in a prototype phase. Another area of prototyping is for the inclusion and use of time-domain metadata in Aladin Desktop and Aladin Lite. This includes work within IVOA to expand on the MOC system that is used for spatial sky coverage, to use the same framework for "time coverage" - a prototype in Aladin Desktop has been shown at IVOA meetings, and we outline the python aspects in Section 4.5.

---

[25] https://aladin.u-strasbg.fr/java/AladinManuelV10.pdf

Aladin Lite, the lightweight version of the Aladin that runs in web browsers, continues to attract a lot of attention with a very large number of external web sites now implementing Aladin Lite as part of their own services. As such Aladin Lite has been one of the drivers of the growth and visibility of the HiPS network, which we attribute to its ease of implementation, and the ability to customise it and integrate it into a wide range of interfaces.

In 2017-2018 we count 350 implementations of Aladin Lite in external web pages, a growth of about 12% compared to the previous year. These range from very simple default implementations to more complex interfaces. The load of queries generated by these implementations continues to be manageable by the system, and we note that the dominant use of Aladin Lite (~70% of launches) is from our own CDS services as Aladin Lite is part of the default SIMBAD results page. It is however very interesting to see the top ten most-used external implementations shown in Appendix B.

A number of improvements have been made to Aladin Lite: HTTPS support has been implemented enabling Aladin Lite to be embedded in HTTPS pages. A SIMBAD pointer feature has been added for quick look-up of SIMBAD sources from the Aladin Lite display. Support for mobile devices has been improved enabling pinch-to-zoom, following high demand for this feature. The main drawing algorithm for the display of the HiPS tiles has also been improved, reducing the amount of flickering when zooming into higher resolution HiPS tiles.

4.4 The CDS Catalogue Cross-Match Service

The X-Match service continues to operate smoothly, with another doubling in the number of queries submitted via the programmatic interface (HTTP API). See the table of X-Match usage below.

A major hardware update for the X-Match service in underway (supported by CNES APR funds). This includes the installation of a set of SSD disks (20 TB) for one of the two servers. We have also made a number of internal modifications to better face bursts of activities like the one following the publication of the Gaia DR2 catalogue. The best long term solution will probably be to offer a programmatic access to the submission of asynchronous jobs. The X-Match service has also been the subject of a number of the R&D topics, see Section 4.6.

CDS X-Match Service Usage

	2013	2014	2015	2016	2017	**2018**
Web interface (jobs/day)	15	16	20	30	33	**43**
HTTP API (jobs/day)	47	50	580	889	1256	**2781**
Associations /day (Web Interface)	~13,000,000	~70,000,000	~55,000,000	~104,000,000	~164,000,000	**196,000,000**
Associations /day (HTTP API)	~298,000	~1,600,000	~6,600,000	~6,700,000	~17,800,000	**53,700,000**

4.5 Service Integration and Portal Development

CDS Portal

One of highlights in the 2017 report was the release of the CDS portal[26]. While the use of the services (SIMBAD, VizieR, Aladin) via the portal is incorporated into the individual service statistics above, we have also monitored the number of unique visitors using the portal, which has stayed relatively stable in 2017-2018 at ~2300 per month, an increase of about 5% compared to its first year of operation.

The portal is designed as a simple entry point to the CDS services, and dedicated efforts were made in 2017-2018 to advertise it. The portal has been demonstrated during a focus demo at the ADASS conference, and was highlighted as one of the new tools shown at CDS booth during the AAS winter meeting in January 2018. Also a number of CDS tutorials have been updated to use the CDS portal as the starting point.

The operation of the portal is in a stable phase. An exploratory development of an interactive spectrum viewer was made using a 1 month contractor in summer 2018. This development has focused on providing an interactive preview of spectra provided by the VizieR associated data service.

Python

An extensive work has been performed to increase CDS engagement in the Python ecosystem. New libraries to access CDS services have been developed, existing libraries have been improved and extended.

A new module, astroquery.cds, has been developed to access the MOC Server, enabling fast retrieval of available datasets in a given region on the sky and filtering on additional metadata. Spatial constraints can be a simple cone, a polygon, or a more complex shape described by a MOC object (an IVOA standard describing spatial coverage). This development has been fully integrated in the popular astroquery package, including documentation[27].

The existing MOCPy library has been extended to support T-MOC (temporal MOC), the prototype counterpart of spatial MOC to characterize temporal coverage of a dataset. This effort will inform development of IVOA standards for the time domain. Available features include read/write of TMOC files, creation of TMOC from a set of time intervals, operations on TMOC (union, intersection, complement) and visualisation as a time barcode (Fig. 7).

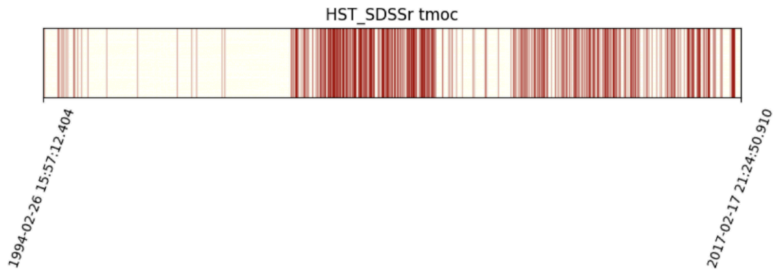

Fig 7.— An example of a Temporal Coverage T-MOC for the SDSS survey

[26] http://cdsportal.u-strasbg.fr

[27] https://astroquery.readthedocs.io/en/latest/cds/cds.html

Other updates have been made to MOCPy: improvement of performance when creating a spatial MOC from a list of positions, faster filtering of a table by a spatial MOC, optimization of memory usage, and a switch to a BSD license made possible by using the astropy_healpix library. Unit tests and continuous integration tools are now in use, ensuring a better stability of the library. A major code restructuring is underway with the purpose of an integration in the Astropy Regions package.

A meeting was held 2-3 July 2018 bringing together HiPS/HEALPix experts and Python developers from CDS and MPIK (Max-Planck-Institut für Kernphysik, Heidelberg) to exchange knowledge, discuss plans and foster collaboration on existing Python packages related to HiPS and HEALPix.

VO datasets browser Javascript widget

We have developed a working prototype of a modular Javascript component, aimed at displaying and accessing available VO data collections, in a similar fashion to the data access tree available in Aladin Desktop. This development should lead to a reusable component, that can be connected to Aladin Lite but is designed to work independently in other contexts and serve as a core component of a VO portal. Next steps include offering this component as a widget for Jupyter notebooks.

HTTPS

The deployment of HTTPS support (a web protocol for secure communication) has been a priority over the past 2 years. Following the initial progress reported in 2017, HTTPS has been deployed on the CDS VizieR servers and on the CDS annotations service. As such, all CDS services can now be accessed through HTTPS.

CDS HEALPix library

HEALPix has become a major key component of CDS products and services. In addition to its use in HiPS, the data stored in SIMBAD and TAP VizieR are indexed according to the HEALPix tesselation scheme. Following these developments we made a specific effort to improve our HEALPix expertise and develop our own implementation of a Java HEALPix library, with a particular focus on accuracy and performance. It is also a strategic decision to develop the local expertise in the fundamental aspects of this by having our own ground-up development and where we are fully in control of the license. The new library is currently being tested and integrated into Aladin Desktop. This Java library will serve as the basis to generate some HEALPix Javascript code to be integrated in Aladin Lite. This will improve HEALPix support in Aladin Lite, bringing the deepest addressable resolution from order 13 to order 24.

We have also made progress with the astropy-healpix package developers to see how our development could help in providing a comprehensive HEALPix implementation for the Astropy community.

4.6 R&D

In the 2017-2018 period we have pursued a varied R&D program with both operational actions and exploratory work to prepare the future. This work was carried out by CDS software engineers with the help of 10 interns and 2 short contracts.

A graphical interface was developed to set options of the COSIM process (the cross identification between Vizier and new catalogues to ingest in SIMBAD). It uses cached SIMBAD results enabling documentalists to more easily optimize the criteria of matching astronomical objects, with the result automatically visualised with Aladin Lite.

An internship was dedicated to the development of a landing-page for VizieR catalogues which is part of the preparation for the integration of Digital Object Identifiers (DOIs) in VizieR. This work is now entering into an operational phase.

A temporal exploration in Aladin was started with experiments involving resource discovery by time constraints, time plots and time coverage.

A short contract was also dedicated to Aladin and the implementation of Astrometry.net

A prototype of a spectrum viewer has been developed by a 1-month contractor. This development has focused on providing an interactive preview of spectra provided by the VizieR associated data service.

An another prototype has been developed with the PWA (Progressive Web Application) technology as a way to provide a version of the CDS Portal which would work better on mobile devices. This work allows the rendering of the CDS Portal client interface to be adapted to the quality of the network bandwidth, as is needed for mobile devices.

In the frame of a long internship we continued the study of Apache Spark technology for its use in the CDS X-Match service. This was done through various collaborations in particular with CNRS-IN2P3. The aim of this study was to simplify the X-Match and to scale it up as easily as possible. This work is a part of a larger study that aims to enable code execution as close as possible to the data. A second and shorter internship in this area of research was dedicated to the test of Cassandra, with a translation of a part of the Java code in Rust and also an experiment with WebAssembly by porting in this language our kd-tree library: with very positive and encouraging results.

We tested the automated classification prototype service to update astrometric posteriors with photometric data and compared our results with an external tool developed by Mara Salvato.

We continued also the work around 3D Visualisation in a Web browser. The objective is to visualise progressively large simulation data or large catalogues whose volume (several TB) exceeds the capacities of a browser. We contributed to a paper written for the SF2A by Frédéric Marin who has used our prototype for multi-scale three-dimensional visualization of emission, scattering and absorption in active galactic nuclei. An another paper is under writing for Astronomy & Computing.

In the frame of IVOA Provenance and SIMBAD Data Models, a design and implementation in RDF/ TripleStore was also studied. A TripleStore database together with a HEALPix spatial index encoding have proven to be very effective and scale efficiently with for the eight million objects in SIMBAD.

We have obtained Research Credits from Amazon AWS to carry out tests with CDS services. The DSS Colour HiPS was made available for a period of 1 month on Amazon Cloud Front as a public HiPS server. HiPS surveys were generated using the computing capacities of Amazon EC2 to evaluate whether there is a gain to doing this on the cloud compared to performing the computations on our own dedicated servers. At this stage, we find that the use of Amazon is more expensive and slower than our local solution. We have also tested Spark / Cassandra for the X-Match and Amazon Aurora for SIMBAD.

An effort was also made around the network and hardware side with an internship concerning the supervision of the Observatory infrastructure.

A “communication and design” intern worked on the definition of a CDS channel to publish video tutorials and also on the social media aspects. He has also produced a video clip about the CDS and a shorter one about the IVOA.

4.7 CDS Infrastructure and Disaster Recovery Plan

In 2017-18 much effort has gone into future planning for the CDS IT infrastructure and for formalising the Disaster Recovery Plan. The new ObAS System Engineer plays the leading role in this effort, and it has already resulted in a number of significant improvements, and a review of the current disaster recovery procedures. The request of the Council in its 2017 report to see the disaster recovery plans has also helped to focus this effort. A number of these improvements are being done at the level of the Observatory as the IT infrastructure is managed within shared server rooms.

CDS IT Infrastructure improvements

Improvements have been made to the network connection and its security. The bandwidth has been upgraded from 1 to 10 Gbit/s, and the redundancy has been improved by an additional link to our internet provider (Osiris). In terms of security, a new firewall has been installed consisting of two new dedicated physical servers running the Stormshield licensed software. The configuration of the network has been changed so that there are now separations between the public access, the internal servers and the internal user network. We have also done the preparatory work for IP isolation, which is an essential part of the network security to avoid internal clashes of user machines with service IP addresses.

A new virtual server infrastructure has been installed. This is designed to support a high level of availability of CDS services, and to simplify the IT infrastructure where small services can be run in a virtual environment rather than on a dedicated hardware server. The new system allows for a new server to be deployed in only 5 minutes, and it provides flexibility in the allocation of resources to a given service. The new virtual server is currently in a test phase where it is being used for non-critical services, running about 20 virtual machines. In 2019 we expect to move into a production phase when we are sure that we can host critical applications and services on it.

Disaster Recovery Plans

The CDS has long had a well established set of procedures for recovering from downtime of the CDS service servers. The organisation of the servers in two servers rooms on the ObAS site, and the back-up of various CDS services on mirror sites is part of the overall security already in place. The recent transfer of responsibility for the ObAS IT infrastructure to the new ObAS system engineer has been a good moment to revise all aspects of the security of our systems. We have made a thorough review of all the existing systems leading to the creation of a catalogue of CDS service components, and we have set up a roadmap for formalising the Disaster Recovery Plan with the aim to be in line with standard practices for a data centre of the size and scope of the CDS.

Following the review of the current status in early 2018 we have decided to invest in the formalisation of the Disaster Recovery Plan. This has involved all of the CDS engineers in discussions lead by the ObAS system engineer. Following meetings in May 2018 and an information gathering exercise, we have compiled a Catalogue of “CDS IT Services” including the shared components with ObAS. This has been completed in summer 2018. We have also sought a formal external training course for the ObAS system engineer to become familiar with industry standards for disaster recovery plans. He was registered for a course in June 2018, but this was delayed and the course was attended very recently on 27-28 September 2018.

The next steps for developing the formal plan are shown below, with the details to be updated from the recent training experience. The major items that need to be integrated into the plan include the role of the Unistra Data Centre, and the timing of the CDS All-Sky Data project.

The table below summarises the current state of the CDS Infrastructure and the various elements that are in place for managing the security of the systems.

Current CDS Infrastructure and Operations Security

Infrastructure	Redundant storage in two local server rooms
	Server redundancy for specific servers
	Back-ups of server configurations, code and data for all services
	Air-conditioned server room with redundant power supply (battery system)
	An inventory application which collects and maintains: - *information on all servers* - *information of what is running on all servers* - *information on the physical location of all servers in the racks (new in 2018)*
	Redundant 10 Gb/s firewalls
	Redundant 10 Gb/s internet connection
Application and Service Code	CDS source code archived with GIT on our CDS GIT server
	Mirrors of CDS services with data and code duplicated for some services (not all services)
	Services designed for redundancy
Monitoring system	Centreon, Nagios and Zabbix supervision servers - *collects information that is presented on service monitoring dash-boards*
	CDS GLU system - long term management for links
Documentation	**The blue-folder "survival kit"** - *organisational schemas of service architectures (currently ad-hoc per service)* - *set of instructions for how to re-boot all CDS services* - *authorisation management for permissions to re-boot services*
	CDS wiki system documentation
	Shared folders

Data Recovery Plan development to be followed in 2018-2019:

1. Identify the critical services and data (IT infrastructure architecture map - replacing our ad-hoc schema)
2. Formalise the actual operation processes
3. Identify risks and provide realistic estimates for recovery times
4. Formalise the Disaster Recovery Plan documentation
5. Identify the improvements necessary
6. Define targets in terms of an internal Service Level Agreement
7. Identify the scope of overall CDS/ObAS security plan for 5 years

5. Projects

Virtual Observatory, ASTERICS

CDS continues to play a leading role in the development of the Virtual Observatory. The full list of contributions to the IVOA are listed in the document provided “CDS Participation in IVOA” which has been updated with a list of the standards developed in 2017-2018, and the CDS contributions to the IVOA interoperability meetings in Santiago, October 2017, and Victoria, May 2018. We gratefully acknowledge the long term collaboration of M. Louys (ICUBE) with CDS on the development of the VO, and we also note the very fruitful local collaboration with L. Michel (ObAS/XMM-SSC). Their contributions to IVOA are included in the document.

CDS staff currently hold a high number of Chair positions in the IVOA. These are listed below. Recent changes include P. Fernique and F. Bonnarel coming to the end of their terms as chairs of the Applications and Data Access Layer Working Groups in May 2018. André Schaaff has become the Chair of the Data curation and Preservation Working Group.

- **Chair of the IVOA Executive Committee** - M. Allen (representing Euro-VO)
- **Executive Board member for France** - F. Genova (representing OV-France)
- **Chair of the Applications Working Group** - P. Fernique *[term ended May 2017]*
- **Chair of the Data Access Layer (DAL) Working Group** - F. Bonnarel *[term ended May 2017]*
- **Chair of the Semantics Working Group** - M. Louys
- **Chair of the Time Domain Interest Group** - A. Nebot
- **Chair of the Data Curation and Preservation Working Group** - A. Schaaff

CDS staff contribute to a wide range of IVOA Working Groups and Interest Groups, and these efforts are aligned with the needs of the CDS in terms of the standardisation that greatly benefits the CDS tools and promotes their widespread use. One of the most active areas is the development of the concepts, standards and tools for supporting Time Domain astronomy. This is one of the scientific priorities of IVOA, and Ada Nebot is providing scientific leadership in this area. CDS interests in time domain astronomy have also been developed in the past year including the extension of MOC to T-MOC, with supporting libraries. Also the time metadata in VizieR for catalogues, but also for time series in the associated data service. Time domain is also a priority in the ASTERICS project, where CDS is interacting with partners such as SKA, CTA, EGO/VIRGO and partners connected to LSST on the requirements for time domain interoperability. Other areas where CDS effort have been prominent in IVOA are in Applications, Data Access, Semantics and Data Models including provenance.

CDS staff have been active in promoting the scientific use of the VO in VO schools, and other events. A special event in 2018 was the triennial IAU General Assembly in Vienna in August 2018, where CDS lead the organisation of the IVOA booth in the exhibit hall. Another special event was the presentation of the IVOA (by Mark Allen), at the invitation of the IAU, in the United Nations COPUOS[28] sessions, in Vienna in February 2018.

The ASTERICS project is now 42 months into its 48 month program. In 2017-18 we have followed the work plan of the project, contributing to the periodic report in 2018 and other planned deliverables. We have also been very reactive to hot topics that have come up, with special focus meetings about ‘time series data’ and ‘time domain’ standards. We have also helped plan the ASTERICS Policy Forum, leading various group discussions during the well attended Policy Forum event.

---

[28] Committee on the Peaceful Uses of Outer Space (COPUOS), http://www.unoosa.org/oosa/en/ourwork/copuos/index.html

Important events for the ASTERICS DADI work in 2017-2018 are shown in the list below, with CDS hosted events shown in bold:

- ADASS XXVII Conference, Santiago, Chile, 22-26 October 2017
- IVOA Interoperability Meeting, Santiago, Chile, 27-29 October 2017
- ASTERICS DADI Virtual Observatory School, Madrid, Spain, 14-16 November 2017
- ASTERICS ESFRI Forum and Training Event 2, Trieste, Italy, 13-14 December 2017
- **DADI Time Series Meeting, Strasbourg, 5-6 December 2017**
- ASTERICS Policy Forum - Steps toward multi-λ, multi-messenger astrophysics, Nice, 17-18 Jan 2018
- ASTERICS "All Hands" meeting, Amsterdam, 14-15 March 2018
- ASTERICS DADI Technology Forum 4, Edinburgh, UK, 16-17 April 2018
- European Data Provider Forum and Training Event 2, Heidelberg, Germany, 27-28 June 2018
- RDA 11th Plenary meting, Berlin, Germany, 21-23 March 2018
- IVOA Interoperability meeting, Victoria, Canada, 28 May - 01 June 2018
- **Strasbourg Time Series Data Model meeting, Strasbourg, 17-20 July 2018**

The CDS will host a number of ASTERICS events before the end of the project in April 2019. The 4th Virtual Observatory School will be held in November 2018, and the Technology Forum 5 will be held in early 2019. We will also host focus meetings related to the new solar physics partner in ASTERICS, the European Solar Telescope.

A major conference is being planned for the end of the ASTERICS project, the "New era of Multi-Messenger Astrophysics" conference[29] to be held in Groningen in March 2019. CDS participants in ASTERICS are playing a role in the organisation, and will also feature on the program, and with a demonstration booth.

ESCAPE

As announced in the highlights, a new project has been successfully proposed in the Horizon 2020 Work Programme INFRAEOSC-04-2018 Call – "Connecting ESFRI infrastructures through Cluster projects". The **ESCAPE** (**E**uropean **S**cience **C**luster of **A**stronomy & **P**article physics **E**SFRI research infrastructures) project will begin in early 2019. The abstract of the ESCAPE proposal is shown below, and a schematic representation of the project is shown in Fig. 8.

> *ESCAPE (European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures) aims to address the Open Science challenges shared by ESFRI facilities (SKA, CTA, KM3Net, EST, ELT, HL-LHC, FAIR) as well as other pan- European research infrastructures (CERN, ESO, JIVE) in astronomy and particle physics. ESCAPE actions will be focused on developing solutions for the large data sets handled by the ESFRI facilities. These solutions shall: i) connect ESFRI projects to EOSC ensuring integration of data and tools; ii) foster common approaches to implement open-data stewardship; iii) establish interoperability within EOSC as an integrated multi-messenger facility for fundamental science. To accomplish these objectives ESCAPE will unite astrophysics and particle physics communities with proven expertise in computing and data management by setting up a data infrastructure beyond the current state-of-the-art in support of the FAIR principles. These joint efforts are expected result into a data-lake infrastructure as cloud open-science analysis facility linked with the EOSC. ESCAPE supports already existing infrastructure*

[29] http://multi-messenger.asterics2020.eu

such as astronomy Virtual Observatory to connect with the EOSC. With the commitment from various ESFRI projects in the cluster, ESCAPE will develop and integrate the EOSC catalogue with a dedicated catalogue of open source analysis software. This catalogue will provide researchers across the disciplines with new software tools and services developed by astronomy and particle physics community. Through this catalogue ESCAPE will strive to cater researchers with consistent access to an integrated open-science platform for data-analysis workflows. As a result, a large community "foundation" approach for cross-fertilisation and continuous development will be strengthened. ESCAPE has the ambition to be a flagship for scientific and societal impact that the EOSC can deliver.

The CDS leads a major work package in the ESCAPE project. The description of this work package from the proposal is as follows:

WP4 CEVO (*Connecting ESFRI projects to EOSC through VO framework*) plans to make the seamless connection of ESFRI and other astronomy and astroparticle research infrastructures to the EOSC through the VO. This implies the need to scale the VO framework to the biggest data sets that will be produced by the ESFRI and other projects. Astronomy has built an operational interoperability infrastructure, the Virtual Observatory (VO) that has proven to be a great success for many aspects of astronomy data interoperability. The VO is an essential component of the astronomy data landscape, as has been strongly stressed in the ASTRONET Infrastructure Roadmap since its first publication in 2008. International astronomy data providers, in particular ground- and space-based telescopes, publish their data using the IVOA standards, and compliant scientific tools and services enable discovery, access and use of the data by the whole astronomy research community.
Key outputs: Assess and implement the connection of the ESFRI and other astronomy and astroparticle RIs to the EOSC through the Virtual Observatory framework, actively contributing to the setting up of the EOSC services.

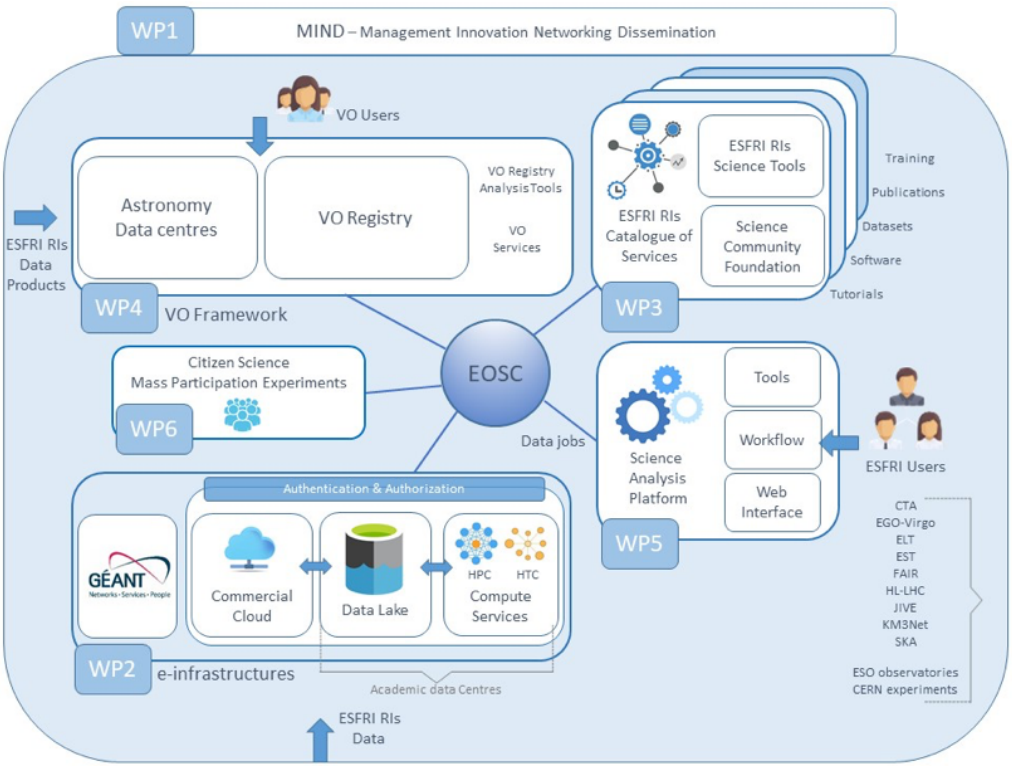

Fig 8. Organisation of the ESCAPE project Work Packages

VESPA / EuroPlanet Project

CDS is participating in the VESPA (Virtual European Solar and Planetary Access) activity in the Europlanet 2020 Research Infrastructure programme funded under Horizon 2020 (2015-2019). While only a small fraction of time is being spent on this activity (1.3 person months in 2017) the impact is high as it bridges CDS astrophysics developments to the planetary science community. Following the steps that we outlined in the 2017 Council report, we note the current status:

- CDS has helped the Paris Data Centre to establish itself as a Planetary HiPS node
- CDS is currently generating planetary HiPS in collaboration with Paris Data Centre, with 45 planetary maps now converted to the multi-resolution HiPS format.
- Progress has been made in Aladin V10 for planetary data support, with the results visible in the public beta version
- Aladin Lite has been adapted to support reverse longitude display for planetary surfaces

As part of VESPA “task 11.4 Planetary Surfaces”, the CDS Aladin Desktop tool has been upgraded to support planetary images and object catalogues, and to allow searches for intersections of complex footprints (CNRS/CDS). A selection of 45 planetary maps from USGS has been converted to multiresolution format (HiPS) and they are available in the Aladin data tree, enabling the plotting of maps quickly using adaptive resolution (see Fig 9.). This also provides a very efficient way to navigate large panoramic images.

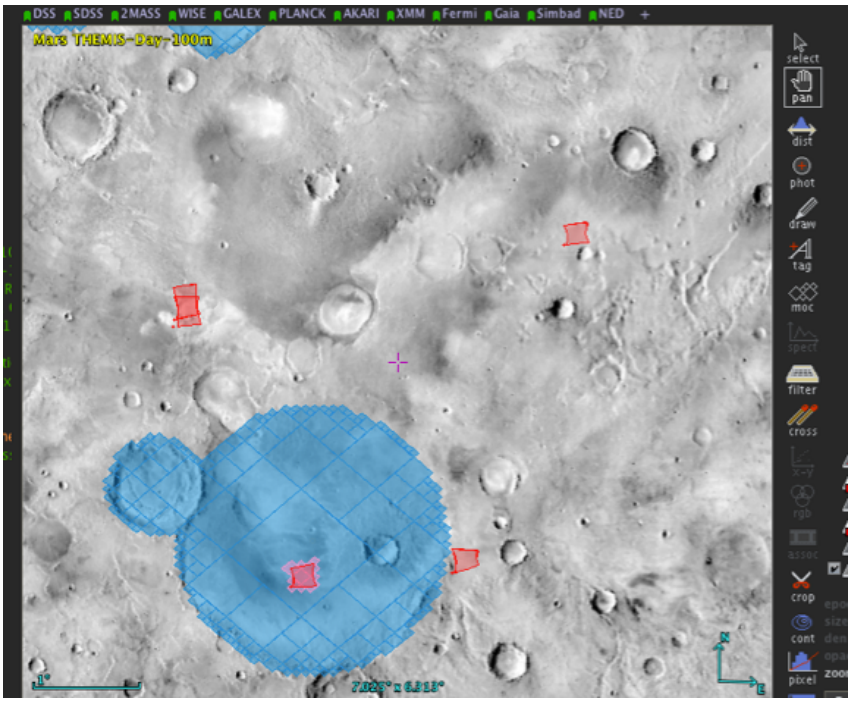

Fig 9. HiPS and MOC in use for managing areas on planetary surface maps

AENEAS

CDS is a minor partner in the AENEAS (Advanced European Network of E-infrastructures for Astronomy with the SKA) project. AENEAS is a 3 year initiative funded by the Horizon 2020 program of the European Commission to develop a science-driven, functional design for a distributed, federated European Science Data Centre (ESDC) to support the astronomical community once the largest radio telescope in the world - the Square Kilometre Array (SKA) -

becomes operational. CDS participation is mainly in Work Package 5, 'Access and Knowledge Creation' which is focused on the interface between a distributed European SKA Data Centre (ESDC) and a distributed body of end users whose goal is the exploitation of SKA data for knowledge creation. WP5 will therefore study the design of "user interaction models" that could be implemented for the ESDC, including the interface with the Virtual observatory.

The Research Data Alliance

CDS has been active in the Research Data Alliance (RDA). F. Genova was re-elected in late 2017 as the co-chair of the Technical Advisory Board (TAB) for a mandate of 2 years. As mentioned in the highlights, RDA Europe 4.0 project started in March 2018. It is set up as a National RDA nodes, with F. Genova is in charge of the French node[30] in collaboration with Françis André from the CNRS Direction of Scientific and Technical Information (DIST). As described in Section 2, support to RDA and RDA-France is mentioned in the National Plan for Open Science.

The French node is the liaison between the RDA and the RDA-Europe project and the national community, and it organizes French participation in the RDA using a model similar to the successful French VO, but on a much wider scale. It obviously fulfils a need: the national mailing list, which had been created in 2014 had 143 participants when it was advertised in the framework of the National Node on 23 May 2018. There are now 337 (6 October 2018). RDA France is progressively getting in touch with the French research organisations, including the CNRS governance, and is setting up a network of correspondents at CNRS. F. Genova and F. André are invited to present the RDA and RDA France in many meetings of different kinds. The connection with Unistra is also well established. The annual meeting of the French node will be held at MESRI on 5 December 2018, in the framework of a high profile event, the "*Journées pour la Science Ouverte*", which will be held at the MESRI 4-6 December 2018. The Journées SOC is chaired by the Open Science Advisor to the Director-General for Research and Innovation of the MESRI.

[30] https://www.rd-alliance.org/groups/rda-france

6. Status of the 2017 Recommendations of the Scientific Council

Here we provide responses to the recommendations that were made in the 2017 report of the CDS Scientific Council. The recommendations are included (in times font) with the original numbering from the report, followed by our response.

3.1 Staffing
The mission of the CDS depends on a core team of dedicated scientists, engineers and documentalists who work closely together. As astronomy moves into the era of big data the future success of CDS depends on improved efficiencies and the innovative use of new technologies, but most of all it depends on continuing to have a good team of the appropriate size and mix of skills. No doubt contributions from postdocs will help, as will the extra efforts that come with support for projects. Nevertheless Council have concerns around the crucial area of permanent staffing and note the following:

- Even with maximizing efficiencies there is a clear need for an expanding workforce as the environment evolves and data volumes grow and complexify
- There will be a critical need for scientific support for SIMBAD around 2019-2020
- More support is required for the VizieR associated data services
- There is a need to accommodate anticipated absences and shifts, as well as to avoid single point failures
- CDS documentalists have specialised roles that need to be recognized and supported for recruitments

3.1.1 Staffing – astronomers
The CDS astronomers play a critical role in the day-to-day quality control of the data curation, the longer term steering of the operations, and the orientation of the research for development. In the context of ever increasing data volumes and of the forthcoming retirement of a key astronomer who works full time for CDS, the council encourages the recruitment of additional astronomers. It understands that these positions are not under the direct control of the authorities, but it expects that judicious use of "coloriage" will attract highly qualified candidates that will make strong positive impressions on the recruitment committees. CDS should also be cognizant of opportunities for attracting research staff from elsewhere in France.

3.1.2 Staffing – documentalist
The Council concurs with the CDS Director's assessment that a documentalist position is needed as a high priority. A documentalist is at the heart of CDS operations. He/she is not just a librarian, but possesses skills needed for data curation. As the HCERES report rightly points out, such individuals need to be trained to support data ingestion from literature with high level of quality and scientific guidance, and an effort needs to be made to keep developing both domain knowledge and technical knowledge. The Council would like to emphasize the specialized nature of the job, which is more than that of a librarian. There is an urgent and immediate need to replace the contract post with a permanent one to deal with volumes of data that are currently being experienced.

After consulting with the relevant parties CDS might consider renaming the documentalist post as, e.g. "data curator", "content curator" or "scientific data curator" with a view to clarifying what they actually do and distinguishing them from librarians.

Response:

We thank the council for this recommendation which reflects the current challenges for the staffing of the CDS. The increase in the volume of journal articles, and the increase in the number of tables and information to be extracted from the journal articles strongly calls for an increase in the staff. The situation is already at a critical level, where the delays on some parts of the CDS ingestion process are becoming unreasonable, and the balance between quality and quantity is recognised by all our staff to be at a very dangerous level. Quality is currently maintained by the exceptional efforts of the staff but the pressure required to continue to achieve this is not sustainable.

We continue to emphasise the specialised roles of the CDS Documentalists who are true ‘stewards’ of the data in the language being used by EOSC. The urgent and immediate need for a replacement of a contract documentalist with a permanent post has been indicated as a high priority.

In terms of scientific support we have consistently alerted our authorities on the critical need for scientific support of SIMBAD on the timescale of 2019-2020, noting the retirement of a senior CDS staff member planned for ~2020. SIMBAD is a core service of the CDS, and a key element of the CDS as a Research Infrastructure. The recruitment of a researcher to provide scientific support for the CDS services is absolutely essential on this timescale.

The long term sustainability of SIMBAD and VizieR requires duplication of the effort, and it should be recognised that these services have expanded in scope in response to scientific needs (e.g. VizieR associated data, complex queries, high volume demands). There is an immediate need to support major works on the core of SIMBAD and to re-new major elements of VizieR. A highly experienced engineer has been hired on a contract (May 2018) to help with these immediate tasks so that operations and development can be managed. A longer term solution requires recruitment of an engineer.

Contractors are an important element of the CDS staff. In the 2017-2018 period CDS has had the following contract staff:

- Six contract staff working on core CDS functions (3 documentalists, 3 Software Engineers). One contract documentalist finished his contract in February 2018. A new contract documentalist, E. Collas was very successfully integrated into the SIMBAD team, starting in November 2017. As mentioned in the VizieR report, T. Pouvreau, hired as a contract documentalist in March 2017, successfully completed her training by the end of 2017.

- Three contract staff on specific projects (ASTERICS - 2 Software Engineers, 1 CNES supported Software Engineer for specific Gaia work)

- Two Postdocs — one in the strategic area of radio astronomy, with time allocated to ingestion of radio astronomy data into CDS services. One ASTERICS postdoc providing scientific support to ASTERICS activities (the previous ASTERICS postdoc, J. Sorce, finished her contract in December 2017)

- One scientific support contract hired to work on ingestion of data into the VizieR associated data service.

- Ten short term interns, and two short term contracts in 2018 as detailed in the ‘Trainees at CDS’ document

3.2 Data Centre at the University of Strasbourg

The University of Strasbourg (Unistra) is constructing a data centre, which could become operational by 2019. Two racks have been reserved for CDS at this data centre. While these are expected be used for a back-up system for CDS as a first step, it is possible that in the future Unistra will expect CDS to have its main equipment at the data centre, citing that the modern facility will be a more robust and secure environment for CDS data operations. At the present time there is no firm information available on the organisation of the data centre, constraints that may be in place for accessing the equipment, costs and other related matters. CDS has a strong operational record over the many years of its operation, and has been providing highly reliable and satisfactory data services to the international astronomical community. The timely ingestion of new data and providing high quality data services on a 24x7 basis is critically dependent on full control of the system and access, and flexibility of operations. These may not be available to the required extent in a data centre that has been developed keeping in mind the requirements of a large and diverse community of users, with its attendant simplifications and mutualisations of the server infrastructure. The Council will need assurances that any new arrangement will provide at least the same level of flexibility and reliability as the current set up. In this regard we note that last year the major CDS services had an average availability at the level of 99.42% (~51 hours/yr downtime).

As things stand we surmise that it will be necessary for CDS to continue to maintain its data centre and related hardware on its own premises, so that full control and flexibility are maintained. However, CDS could very usefully develop a duplicate facility at the Unistra data centre, which will be fully operational and serve as an essential backup. CDS can also use its long experience to guide developments at the Unistra data centre, and it might therefore be useful if a Unistra representative, with responsibility for the data centre, were able to meet the Council at its next session.

Response:

In 2017-18 we have made efforts to gather information about the Unistra Data Centre and to understand the potential and risks for its use in future CDS operations. A visit of a Unistra Data Centre representative (P. Gris) to ObAS took place in January 2018. He made a general presentation of the data centre, showing it to be in the early phases of construction of the building on the Unistra campus. The presentation outlined the scale of the centre — it has space for 110 racks. The different services to be offered were described. We have been reassured that one of the operational modes is to host physical servers, as would be necessary for managing CDS servers. We continue to develop our plans for a step-wise duplication of CDS servers in the Unistra Data Centre when it becomes operational, with a long transition phase where servers on the ObAS site remain as the master site. The timescale for opening of the centre is early 2020. No information on costs has yet been provided. A representative from the Unistra Data Centre (R. David) will be present at the Scientific Council meeting.

3.3 CDS and Big Data

The Council fully supports the CDS focus on developing its strategy in the era of Big Data, and its strong commitment to maintaining its leadership role as an international reference data centre and service provider. The coming decade will see major data releases (DRs) from missions such as Gaia (several over the next decade), Euclid, VISTA Phase 2, LSST, and SKA. Services that build on interoperability and synergy between these very large datasets will play a crucial role in enabling science exploitation. CDS is well placed to exploit this potential both in terms of its knowhow and as trusted partner. We therefore urge the CDS to devote effort in proactively investigating ways of ensuring early involvement in big data projects, such as Euclid and LSST, in order to be able to exploit these opportunities and also as a means to investigate ways that they might bring added value to these projects. Specifically we encourage the CDS to investigate the potential of providing services internally to consortia as a means for consortia partners to work with proprietary data, though this should not involve providing elements of any pipeline processing architecture. One advantage of this approach is that it will greatly speed up and facilitate open access to the data, once restrictions are lifted. The potential of project partnerships as opportunities for new and additional funding streams should be carefully evaluated by CDS.

Response:

We thank the Council for this recommendation. It is clear that engagement with large projects is an important aspect of the future planning of the activities of the CDS. The highlighted example of the CDS participation in Gaia DPAC shows the benefits of this, as does our continuing collaborations with ESO, ESA, NASA and CNES. The collaboration with SkyMAPPER, and with MAST for PanSTARRS also demonstrates excellent results for making science ready data available at the "data release" phase of these projects. We are also increasingly involved with large projects via our participation in ASTERICS, AENEAS and ESCAPE where we work together with ESFRI projects in the context of the Virtual Observatory (and now also EOSC). CDS has had important interactions with the LSST project within IVOA and at the ADASS conference, in addition to CDS participation in the LSST@Europe3 meeting in June 2018. LSST has adopted various technologies developed by CDS (in particular HiPS and MOC), and these are continuously discussed at the scientific and technical levels. Early involvement in projects with closed partnership agreements and data rights (including LSST) is however a challenge. We have approached some projects (e.g. CFIS) to discuss the use of authenticated services. The establishment of more formal connections to LSST would require guidance from CNRS-INSU. We anticipate a continuing discussion with the Council on these issues.

3.4 The Scientific Council

The members of the current Council were appointed from 2016 to 2018 and we therefore anticipate that the Council will be renewed after the next meeting. The Council were somewhat concerned about the number of their members who were able to attend the meeting and noted that the absence from the closed discussion of all but one of the French Council members, significantly skewing the composition from its nominal 50% French plus 50% international. In view of the fact that there were no controversial issues to discuss in 2017 it was not considered seriously problematic for that meeting. Furthermore, all of the Council will have an opportunity to view and comment on this report. Council request the CDS Director to set up a poll before the end of 2017 with a view to finding a suitable date for the 2018 meeting, that would allow attendance by the maximum number of Council members.

For future meetings it would be helpful to the Council if everyone present at the meetings, i.e. the CDS staff and the Council members, would wear badges indicating their names and positions.

Response:

Following an early poll we were able to find a date which enables a high level attendance of the Council for the 2018 meeting. Badges will be provided at the 2018 meeting.

The three year term of this Council comes to an end in 2018, and a new Council will need to be formed by our authorities, the CNRS-INSU and the Université de Strasbourg.

3.5 Other Matters

The Council would be interested to see (1) a disaster recovery plan and (2) a SWOT analysis for CDS, the absence of which was noted in the HCERES report, but which we understand does exist.

Response:

(1) This report has outlined the current status of the Disaster Recovery Plan, and the engagement of the new ObAS system engineer to lead the formalisation of the plan taking into account the major developments including the CDS All-Sky Data project and the future availability of the Unistra Data Centre.
(2) An informal SWOT analysis did help guide the writing of the HCERES evaluation documents in 2016, but there is no documentation of this SWOT. Apologies for any confusion about that. An intention to perform a new SWOT in 2017-18 was not carried out.

Appendix A. — Table of large HiPS data sets 2017-18

Large HiPS Data sets (2017-2018)

Survey		FITS	JPEG or PNG	Pixels	Date published	HiPS generation
PanSTARRS	DR1 band g	✓	✓	24 Tpix	Dec 2017	CDS
	DR1 band z	✓		24 Tpix	TBD	CDS
	Colour composition		✓	12 Tpix	April 2018	CDS
ESO Surveys	VISTA VIKING H	✓	✓	800 Gpix	April 2018	CDS
	VISTA VIKING J	✓	✓	800 Gpix	July 2018	CDS
	VISTA VIKING K	✓	✓	800 Gpix	July 2018	CDS
	VISTA VIKING Y	✓	✓	800 Gpix	Sep 2018	CDS
	VISTA VIKING Z	✓	✓	800 Gpix	Sep 2018	CDS
	VISTA VVV H	✓	✓	200 Gpix	Sep 2017	CDS
	VISTA VVV J	✓	✓	450 Gpix	Sep 2017	CDS
	VISTA VVV Z	✓	✓	450 Gpix	Sep 2017	CDS
	VISTA VVV Y	✓	✓	450 Gpix	Oct 2017	CDS
DECam Surveys	DECaLS DR5 colour		✓	4 Tpix	June 2018	CDS
	DECaPS colour		✓	1 Tpix	Jan 2018	CDS
	DECaPS g band	✓	✓	6 Tpix	Jan 2018	CDS
SkyMAPPER	6 bands: G,I,R,U,V,Z	✓	✓	6 x 1 Tpix	Dec 2017	CDS
	Colour composition		✓	1 Tpix	Dec 2017	CDS
MAMA	Colour composition		✓	1 Tpix	Dec 2017	CDS
SWIFT	22 bands	✓	✓	22 x 50 Gpix	Oct 2017	HEASARC
Beijing-Arizona Sky Survey (BASS)	DR1 Colour		✓	500 Gpix	Oct 2017	China-VO
	DR2 Colour		✓	600 Gpix	Dec 2017	China-VO
Dark Energy Survey	DR1 Colour		✓	2 Tpix	Dec 2017	NOAO
HERSCHEL	PACS Colour		✓	100 Gpix	Sep 2017	ESA

Appendix B — Table of top ten Aladin Lite implementations by usage

Site Name	URL	Aladin Lite Launches / yr
ESA Sky	http://sky.esa.int/	79668
HyperLeda - database for physics of galaxies	leda.univ-lyon1.fr	72150
The Extragalactic Distance Database	edd.ifa.hawaii.edu	61314
Canadian Astronomy Data Centre (CADC)	http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/AdvancedSearch/	45778
SkyMAPPER	http://skymapper.anu.edu.au/sky-viewer/b(45384)	45384
Cambridge Gaia Alerts	http://gsaweb.ast.cam.ac.uk/alerts/allsky	43501
ESO Archive	archive.eso.org	38062
ALMA archive at NRAO	almascience.nrao.edu	34717
ALMA archive at ESO	almascience.eso.org	34469
Beijing-Arizona Sky Survey (BASS)	http://hips.china-vo.org/bass-dr2-image/index.html	27366