

# Seeking CoreTrustSeal certification to check and expose trustworthiness outside the community boundaries



Strasbourg astronomical Data Centre CDS (Centre de Données astronomique de Strasbourg) adopts Repository Audit and Certification Catalogue RDA Recommendation

*“As often in RDA, the output addresses both technical and sociological challenges. On the technical side, **it allows applicants to check their processes and the technical framework of the repository**, which can lead to identify domains in which quality can be improved. The initial self-evaluation performed to fill the form is a very useful exercise already, and the evaluators may detect additional issues and suggest improvements.*

*Sociological challenges are, as often, the toughest. Seeking external certification was a way to check and expose CDS trustworthiness outside the community boundaries. It is also critically important in a context in which data management plans are more and more required by funders. It is a good point to deposit data in a certified repository. There is a*

*risk that generic repositories, which have good, recognized processes, enter in competition with disciplinary centres. These latter provide a higher level of data stewardship because they have the necessary disciplinary knowledge, and the shift to generic repositories would be at the expense of data and metadata quality and relevance to science needs. It is thus essential to demonstrate the trustworthiness of disciplinary centres. **CTS criteria adequately question the repository expertise, including scientific knowledge and guidance, with respect to its mission.** They are achievable by well-tended small/medium size repositories for which seeking ISO certification would be too heavy and resource-consuming.”*

**Françoise Genova**, Senior Researcher, CDS, Observatoire Astronomique de Strasbourg, France

## The Challenge

Since its creation in 1972, Strasbourg astronomical Data Centre CDS has been developing value-added services which are widely used by the international astronomical community in its daily research work. The CDS hosts:

- » SIMBAD<sup>1</sup>, the world reference database for the identification of astronomical objects;
- » VizieR<sup>2</sup>, the service allowing to browse the CDS reference collection of astronomical catalogues, tables and other data published in academic journals and to extract information from it;
- » the Aladin<sup>3</sup> interactive software sky atlas for access, visualization and analysis of astronomical images, surveys, catalogues, databases and related data.

There were around one million queries per day on the CDS services in 2018 from all around the world. These numbers, as well as the long list of enduring collaborations between CDS and major observatories, space agencies and journals, demonstrate that the astronomical community perceive CDS as trustworthy. They are used as indicators of CDS impact in the discussions with the CDS authorities, and can also be used in general to demonstrate CDS value and its trustworthiness. It was however worth seeking to obtain an external recognition from outside the disciplinary context.

## The RDA outputs adopted

CDS joined the WDS in 2012, soon after it was created. WDS was initially constituted from data providers from the Earth Science and astronomy communities, and was then a natural host for CDS certification. It was also well fit because it fully considered data services: the CDS core mandate is to provide services to the science community, and being a sustainable data repository is somehow a collateral consequence of the provision of value-added services on the long term. In parallel, the Data Seal of Approval was recognized as the first ‘basic’ certification step in the European context. CDS obtained the Seal in 2014. Applying for the CoreTrustSeal, based on the **RDA Repository Audit and Certification Catalogues** recommendation, which merges DSA and WDS requirements and procedures, was then a natural step when it was time to renew the DSA.



Find out more at:

[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit

[rd-alliance.org](http://rd-alliance.org)

or write us at

[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)

## RDA Added Value for CDS

The existence of two 'basic' certification frameworks, DSA and WDS, was really confusing. They initially addressed different science fields but were in fact not specific to any. They had some overlap, but each also brought different points of view and additional criteria which were complementary. It was then worth to apply for both, as the CDS had done. A handful of data repositories and services had gone through the same path, but it was painful to go through the process twice with different requirements to fulfil. In fact, for most repositories one regrettable consequence of confusion was to discourage the repository to apply to any. This was a real issue with the **development of Open Science, in which trustworthiness is critical for adoption by**

**data providers and end users:** this strongly pleads for repositories to go through a formal certification process. 'Basic' certification provides a do-able but rigorous accession path. It was thus really meaningful to merge DSA and WDS.

Members of the DSA and WDS Boards participated from the first Plenary meeting, and discussions during this meeting led to the creation of the RDA/WDS Certification of Digital Repositories IG that defined a programme of work for the RDA Repository Audit and Certification DSA-WDS Partnership WG to merge the two frameworks. **This was an early demonstration of the RDA value as a neutral, global forum, and of its capacity to develop very meaningful recommendations**

## The benefits of adopting the RDA output

The astronomical community has been at the forefront of scientific data sharing, and the scientific community routinely uses on-line services. Ground and space-based observatories provide their data in their archives, value-added services such as the CDS ones are resources for the daily research work, and articles from academic journals are linked with observations and databases. In addition, the Virtual Observatory provides a global, open interoperability framework for data and tools. The online services constitute a global data infrastructure with seamless access to data. The CDS is evaluated every 5 years together with its host, Strasbourg Astronomical Observatory, and in addition every time the National Research Infrastructure Roadmap is updated. Seeking external certification was a way to validate and expose CDS trustworthiness outside the community boundaries.

## The adoption process

Certification was sought for one of the services, Vizier, which has a specific preservation mandate for some of its data holdings.

1. The IT engineer responsible for Vizier took the lead to make a first pass through the criteria. The data steward team was involved in the update of the end-to-end description following the Open Archival Information System (OAIS) model to describe the different activities performed by CDS and their relationship internally and with the external world<sup>4</sup>.
2. The form was then sent to the CDS current and past directors for input on the criteria relevant to the CDS mission, organisation and context, and those requiring a scientist's point of view.
3. Strasbourg Astronomical Observatory System engineer was involved in the answers provided in particular to the questions dealing with security, risk management, and the technical environment (Strasbourg Astronomical Observatory hosts the CDS).
4. The application requested 3-4 Full Time Equivalent weeks. The evaluation was very positive, with only minor comments. It appears that the provision of detailed answers to the criteria following closely the CoreTrustSeal Extended Guidance document was well received<sup>5</sup>.

## Main Takeaways

It is essential to involve the team in the process, or at least the people who have full knowledge of the different tasks and how they interact with each other, and of the different aspects covered by the criteria. The form cannot be filled only by the director, or the technical director, or researchers, or IT engineers, or data stewards. People with all these profiles have to bring their knowledge and input together.

Although it is not formally required, developing an end-to-end description of the processes with the OAIS model in

Find out more at:

[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit

[rd-alliance.org](http://rd-alliance.org)

or write us at

[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)

mind is a useful exercise. It is particularly worthwhile to involve all the people who participate in the processes: they bring their deep knowledge of their tasks and interactions within and outside the repository, and it is also a very good occasion to value each one's contribution.

The Extended Guidance provided by CTS, which was developed after the RDA recommendation completion by DSA and WDS and is now managed by CoreTrustSeal, is a very useful addition to understand which information is relevant for each criteria. There is still some redundancy between criteria, which remains manageable thanks to the guidance.

### CDS recommends adoption without reservation

One of the very interesting aspects of applying for core certification is that it allows repositories to check their own processes. The answers provided to the criteria are checked by external evaluators when the application is submitted, but filling the form already enables the repository to perform a first self-evaluation, to identify issues and to take corrective measures.

Evaluation of the repository trustworthiness by an external, recognized body is an important asset. It can be put forward to the repository authorities and users. Depositing data in a certified repository is a positive point in data management plans and certification will encourage data provider to deposit data in the repository.

It can be noted that CDS encouraged other astronomical data centres to apply to WDS when it became part of it, and data centres from Canada, China, Italy and Russia became members. They never went through DSA certification. They now have vocation to adopt the RDA recommendation by becoming CTS members when their WDS certification is due for renewal. The Chinese Astronomical Data Center of the National Astronomical Observatories, Chinese Academy of Sciences, has been the first astronomical data centre to obtain CTS certification.

## About Strasbourg astronomical Data Centre CDS

Strasbourg astronomical Data Centre CDS (Centre de Données astronomique de Strasbourg) is dedicated to the collection and world-wide distribution of astronomical data and related information. Since its creation in 1972 as the Stellar Data Centre (the name was changed to Strasbourg astronomical Data Centre in 1983), the CDS has been a pioneer in the dissemination of added-value digital scientific data. It has been playing a major role in the international collaboration which defines the astronomical Virtual Observatory (VO). The VO is the global disciplinary interoperability framework which enables seamless access to the wealth of astronomical online resources, which is now implemented by most major astronomical data providers. CDS services, and the Virtual Observatory, are widely used by astronomers world-wide in their daily research work.

### CDS user community is the global astronomical community.

Researchers use the services in their daily research work, and provide data attached to their research articles for inclusion in Vizier collection. CDS collaborates with major agencies and academic journals active in the field, which are also using its services and providing data. The services are also used in astronomy higher education courses, and by the general public interested in astronomy, in

particular amateur astronomers.

In addition, CDS is one of the Research Infrastructures of the National Research Infrastructure Roadmap, which links it to the policy level – information about eventual certification is requested in the applications to the national roadmap. It is also involved in the RDA France National Node, which has chosen Certification as a priority topic, and the experience gained from going through the certification process has already been, and will continue to be used in dissemination activities at the national level (e.g. translation of the CTS guidance, presentations, discussions, specific workshops).

### CDS and RDA

CDS had been involved in the construction of the RDA as a participant, on behalf of the CNRS, in the three first European projects funded in support of RDA, and has thus been well informed about its activities from the start.

**Contact person: Françoise Genova**, Senior Researcher, CDS, Observatoire Astronomique de Strasbourg,



1 <http://simbad.u-strasbg.fr/simbad/>

2 <https://vizier.u-strasbg.fr/viz-bin/VizieR>

3 <https://aladin.u-strasbg.fr>

4 <http://cds.u-strasbg.fr/vizier-org/>

5 <https://www.coretrustseal.org/wp-content/uploads/2019/02/Strasbourg-Astronomical-Data-Centre.pdf>

Find out more at:

[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit

[rd-alliance.org](http://rd-alliance.org)

or write us at

[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)