SF2A 2025
A. Siebert, K. Baillié, M. Béthermin, F. Cantalloube, E. Josselin, N. Lagarde, J. Malzac, J. Richard, L. Selliez, O. Venot (eds)

PARIS ASTRONOMICAL DATA CENTRE: 20 YEARS AND COUNTING

S. Erard[1], P. Le Sidaner[2], C. Boisson[3] and Z. Meliani[3]

Abstract. The Paris Astronomical Data Centre (PADC) celebrates its 20th anniversary this year, marking two decades of commitment to astronomical data services and interoperability. PADC primarily supports local ANO5 services. Its expertise extends to uncommon fields of the VO including the Solar System, modeling, atomic and molecular physics, fostering interdisciplinary interoperability. Beyond its core mission, PADC has played a proactive role in the development of the Virtual Observatory, with significant contributions to Open Science and FAIR principles.

Keywords: Astronomy, Virtual Observatory, Open Science, Interoperability, Data

1 History

The Paris Astronomical Data Centre (PADC) was created in the mid-2000s to provide a stable infrastructure for VO projects and to help integrate data resources into international frameworks. This initially covered projects from a number of institutes in the Paris area. With expanding structuration at national level, this was later restrained to CNRS-certified "ANO5" services at Paris Observatory, starting with VAMDC, VOSSP, MIS & Jets, BASS2000, and the Encyclopedia of exoplanets. Over time, additional services were developed and certified (APIS, VESPA, Plus-Sismique, and MASER) and others were reorganized (`https://padc.obspm.fr/`)

PADC later received a complementary mission to support Open Science activities and to expand interoperability and the application of FAIR principles to a variety of projects at Paris Observatory. PADC therefore contributes to distributing patrimonial and long-tail databases, which complement archives from ground observatories and space agencies. After twenty years, PADC has grown into a transversal activity at the Observatory, covering multiple domains while maintaining its core mission: open, sustainable, and interoperable data access.

2 Functioning

PADC provides and maintains a mutualized infrastructure available to certified services distributing data. This includes physical servers, storage and computational resources, located on the three sites of the Observatory: Paris, Meudon and Nançay. Beyond hardware, PADC also provides and maintains a variety of software tools and supports low level functionalities related to service interoperability (ingestion, registration, monitoring, etc).

A biannual call is open to all Observatory personnel, who can request specific hardware installation, travel funding (typically to Interop meetings of the IVOA and other consortia), methodological support related to interoperability and data access, small development projects, and support for local workshops in this area. Although certified services have priority, any project distributing data in an interoperable manner is eligible.

More generally, PADC is the transverse structure in the institute to acquire and share VO and interoperability expertise between local projects. The guiding principles are flexibility, sustainability, reusability of solutions, interoperability across disciplines, and close collaboration between researchers and engineers, inherited from space instrumentation activity.

[1] LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CY Cergy Paris Université, CNRS, 92190 Meudon, France
[2] DIO, Observatoire de Paris, Université PSL, 75014 Paris
[3] LUX, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, 92190 Meudon

3 Available resources

Most services now use a shared hardware infrastructure that relieves teams of management issues and reduces global and environmental costs. This infrastructure includes two redundant data servers (Meudon and Paris) as well as three hypervisors hosting virtual machines and software servers, tape-based robotics for data archiving, computing servers with PADC priority on the computing cluster (Tycho) and a local cloud (OpenStack), a server equipped with GPU cards for machine learning-type processing. The Nançay Data Center, which is increasingly used to provide access to local data, has been connected to PADC since 2020. The overall storage related to PADC is on the order of 700 TB.

Altogether, PADC maintains approximately 50 virtual machines for data distribution (DaCHS and das2 servers that support access protocols), access portals (VESPA, VAMDC, ObsCore, Exoplanets, CTA, etc.) and on-demand computing (OPUS). In addition, PADC provides various tools such as a GitLab with AAI, monitoring tools, and the UWS online code platform. For published services, PADC offers monitoring, backup, maintenance, and usage statistics.

PADC has no dedicated staff. Instead, it relies on transverse manpower provided by the Observatory's laboratories (Paris-Meudon and Nançay sites) and services (Direction Informatique, the IT service), mostly engineers. This distributed model provides flexibility but requires continuous engagement from the labs to ensure long-term sustainability. The growing shortage of manpower in the academic field however requires to manage services globally with related technics whenever possible, hence to apply standards and uniform procedures.

4 Activity

Beyond certified national data services, PADC also supports a number of community projects as well as primary functions of the Virtual Observatory infrastructure.

PADC contributes to Open Science efforts with DMPs, DIOs, and OS policies at the Observatory of Paris. An important aspect is to provide off-the-shelf solutions for projects under contracts (such as ANR, ERC, etc) to make their data locally available at the end of the contract. This is both economical and gratifying with respect to commercial alternatives. Other projects imply the integration of heterogeneous elements from various fields, e.g. the recent CNES / INSU Data and Service Centers in planetary science which require interfaces with existing certified services and similar projects abroad.

Some development is possible on request, in particular to increase tool connectivity, e.g. to add SAMP and fits support plugins to open source applications. A historical application at PADC is the development of portal frameworks, which have been incrementally reused in several domains: VAMDC, Exoplanets, CTA, VESPA, and the recent ObsCore portal (`https://voparis-portal.obspm.fr/`). This enables small teams to reach a professional level of service with reduced efforts (Fig. 1).

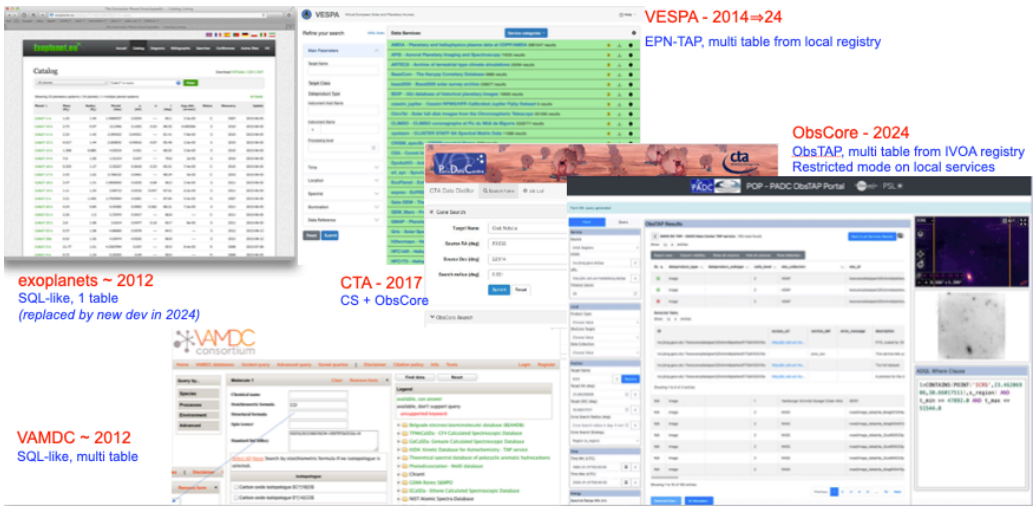

Fig. 1. Portals developed for successive services

The ObsCore portal is the first VO client (with the VESPA portal for Solar System data) to query all data services in the registry, allowing the discovery of unknown datasets and cross-searches between servers. Several local data collections are being published with ObsCore, e.g., images from various telescopes at OHP (an educational project at undergraduate and master level), digitized plates from NAROO, patrimonial databases existing in all teams, and several Pro-Am collaborations (e.g., stellar occultations by asteroids, comet observations, etc).

Primary VO supports also includes mirrors of the IVOA registry and HiPS surveys, VOTable and service validators, monitoring of VO services worldwide, and contributions to community python libraries. Over time, PADC got involved in the large scale VO projects such as the computation grid, workflow platforms, AI, and now the use of EOSC (European Open Science Cloud).

Technology watch consists in assessment of standards and tools, which often lead to applications in unexpected situations - this has been particularly true for Solar System applications of Aladin, TOPCAT, HiPS, MOC, etc. Strong interactions with other VO centres (CDS, GSO, IDOC in particular) help assess tools and ensure interoperability across the data ecosystem.

5 International actions

PADC provides coordination and visibility for international participation of local staff, both to acquire expertise and discuss new standards. The historical and main focus has been the participation to the IVOA Interop meetings, with a large involvement at coordination level (currently four chairs and one co-chair of WG and IG) and with the edition of VO standards (at least seven).

Recently certified services have driven participation to other consortia, in particular to the IPDA (space agencies for space mission data), IHDEA (heliophysics), RDA, and thematic consortia such as VAMDC and Europlanet.

PADC is more and more involved in European programmes and bids, e.g., ACME, EXTRACT, FAIR-IMPACT, OSCARS' Astronomy Competence Centre, and OSTrails programs. The Observatory of Paris is a member of the EOSC association, and PADC is its executive body in this context.

6 Prospects

PADC faces interesting challenges for the coming years:

- Sustaining manpower and lab involvement without dedicated staff.
- Continuing implementation workshops, pioneered in VESPA and now extended to the Astronomy Competence Centre.
- Enforcing the local knowledge about interoperability solutions and the Open Science policy, through educative actions.
- Maintaining excellence in the frame of the evolving policy related to certified services at national level.
- Identifying a sustainable solution to assign DOI to datasets, ensuring compatibility with the national recherche.data.gouv platform.
- Supporting coming ground-based instruments which will provide huge datasets, e.g., CTA, NenuFAR, or SKA.
- Enhancing European/global integration via EC programmes and EOSC.
- Exploring new perspectives, e.g., the ability to rely on VO techniques and tools to manage ground-based and space borne instruments.

PADC is founded by the Observatory of Paris and CNRS / INSU. The author thanks the many colleagues, engineers, and collaborators who contributed to PADC services over the last twenty years, and the organizers of the SF2A 2025!