SF2A 2023
M. N'Diaye, A. Siebert, N. Lagarde, O. Venot, K. Baillié, M. Béthermin, E. Lagadec, J. Malzac, J. Richard (eds)

OPEN SOLAR DATA, DATA PRODUCTS, AND TOOLS AT MEDOC

E. Buchlin[1], S. Caminade[1], S. Parenti[1], B. Perri[1,2], F. Auchère[1], N. Traoré[3], K. Ashkar[1] and N. Dufourg[4]

Abstract. MEDOC (Multi-Experiment Data and Operation Centre), initially created as a European data and operation centre for the SoHO mission, has grown with data from other solar physics space missions, from STEREO to SDO, and now Solar Orbiter. In addition to observational data, MEDOC also provides data products derived from observations (maps, catalogues...), tools for data analysis and interpretation, and numerical simulation results. We present the current and future MEDOC interfaces, including Application Programming Interfaces (APIs) and Virtual Observatory (VO) services, data formats, implementation of Digital Object Identifier (DOIs), and how they contribute making MEDOC data FAIR (Findable, Accessible, Interoperable, and Reusable).

Keywords: open data, archives, interfaces, tools, heliophysics

1 Introduction

The main scientific objectives of solar physics concern the physical mechanisms behind the Sun's internal structure, its magnetism, its atmosphere (in particular the corona) and its wind, as well as their variability. In addition, scientific understanding of these phenomena has applications for space weather, which seeks to understand how they can affect the Earth and human activities, and to anticipate them in order to mitigate their adverse effects.

The complexity of these phenomena, involving many physical fields and many temporal and spatial scales, requires the coordinated use of numerous datasets, covering long periods at high cadence, including measurements of different physical quantities, measurements covering several wavelength ranges, wide field of view observations and others at high resolution. Scientific data centres have an essential role to play in the data life cycle, by allowing the scientific community making the most of these data and exploiting them for optimal scientific return, thanks to their adhesion to the FAIR principles* for Findable, Accessible, Interoperable, and Reusable data and metadata.

MEDOC† (Multi-Experiment Data and Operations Centre) is a data and operations centre for space solar physics, set up in 1996 at the Institut d'Astrophysique Spatiale (IAS) by CNES and CNRS/INSU with the support of the Université Paris-Sud (now Université Paris-Saclay), for the ESA/NASA SoHO mission. Since then, it has received data from other missions and developed numerous value-added products and services. MEDOC's objective is to serve the solar physics and heliosphere scientific community by supplying data, tools, and services, and carrying out scientific operations for space missions in the discipline. MEDOC works in collaboration with the other French data centers and observations services (CNRS/INSU AA-ANO5 and/or ANO6) in heliophysics: CDPP (in-situ measurements), 3SOLEIL (ground-based observations), APIS, MASER, CLIMSO, and STORMS.

[1] Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France
[2] Dept. d'Astrophysique/AIM, CEA/IRFU, CNRS/INSU, Université Paris et Paris-Saclay, 91191 Gif-sur-Yvette Cedex, France
[3] DSI Group, Paris, France
[4] CNES, Toulouse

*https://www.go-fair.org/fair-principles/
†http://medoc.ias.u-psud.fr

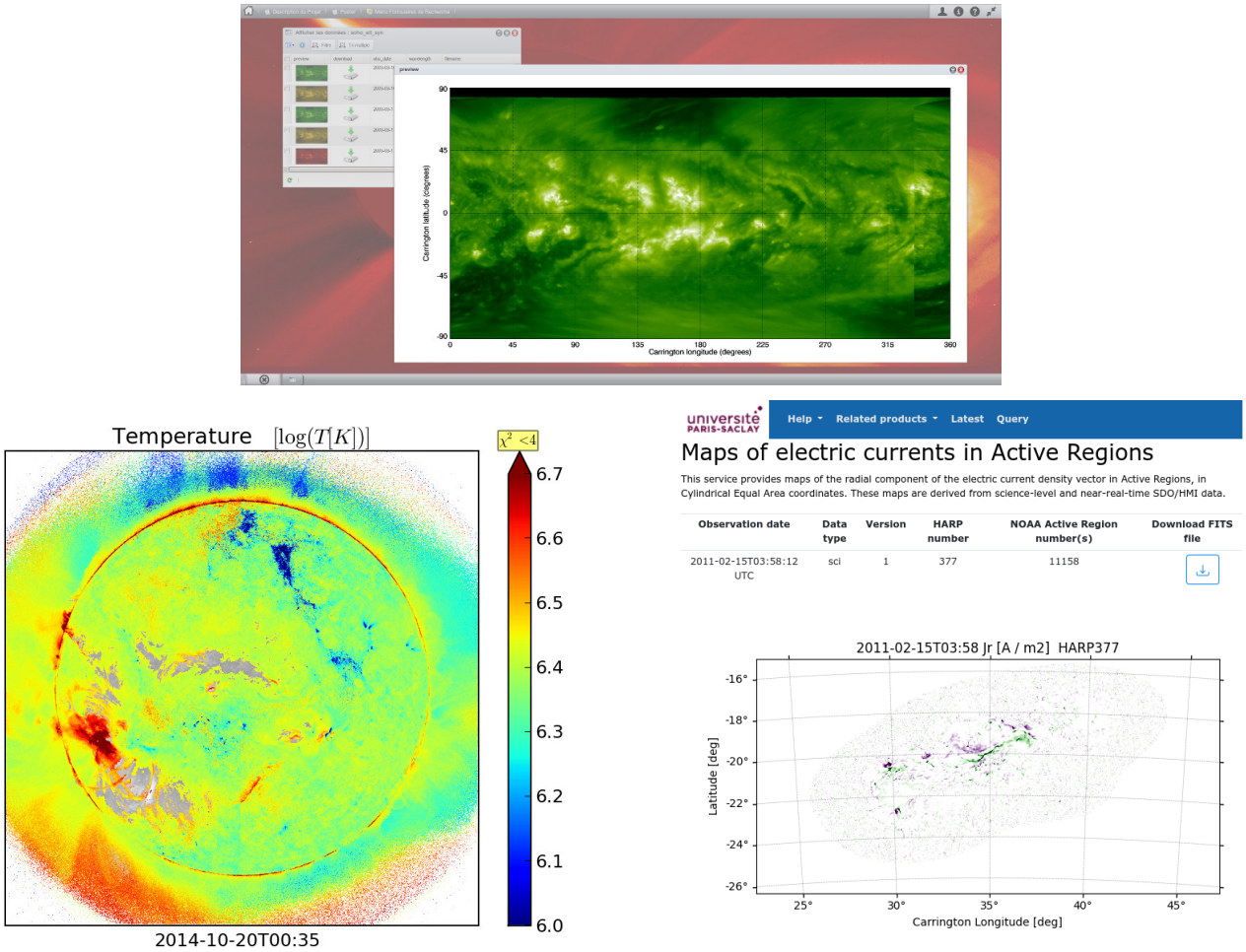

Fig. 1. Examples of derived data products at MEDOC: SoHO/EIT synchronous synoptic maps (top), temperature map of the corona (bottom left), and electric currents in active regions (bottom right). The latter is one of the four products provided by MEDOC to the ESA Space Weather portal.

2 MEDOC observation data and derived data products

MEDOC hosts data from SoHO (1995–), STEREO/SECCHI (2006–), SDO (2010–), TRACE (1998–2010) and other missions, for a total of more than 900 TB. It is also the main mission archive for the CNES PICARD mission (2010–2014), and will include Solar Orbiter data (2020–). From SoHO (1995), almost all solar physics data from space missions are open, with a proprietary period of no more than 3 to 6 months (if any) after telemetry reception, and this proved to be very useful for multi-instrument studies and collaboration. In the heliophysics community, reluctance to this policy, due either to fears about being scooped or to concerns about data quality, exists but is limited.

MEDOC also provides value-added, derived data products, that are computed from observational data stored at MEDOC, such as movies, spectral atlases, catalogs, maps of the thermal structure of the corona and of electric currents in active regions, synchronous synoptic maps of the photosphere and corona (see Fig. 1). Some of these are computed several times per day, just a few hours after the observation itself.

3 Interfaces for data access

The main web interface to data and derived data products is provided by the SiTools2 software developed for CNES (2010–2017) by the AKKA company, based on Java and ExtJS technologies (see Fig. 2). We have developed IDL and Python clients for the SiTools2 REpresentational State Transfer (REST) API, that can be used to analyse large datasets by scripts.

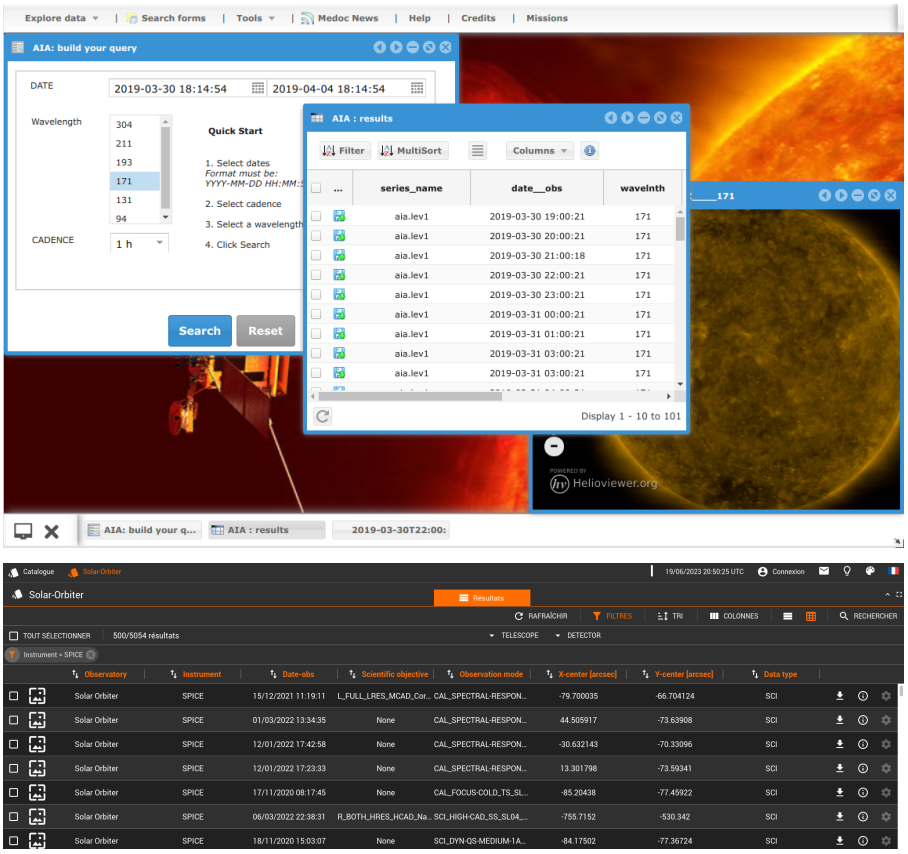

Fig. 2. Top: SiTools2 interface at MEDOC. Bottom: Solar Orbiter data in the current status of the REGARDS interface at MEDOC.

Since 2017, the CS company is developing REGARDS for CNES, as a successor of both SiTools2 and SIPAD (used in particular at CDPP). REGARDS has been installed at MEDOC, and several new functionalities had to be developed to adapt it to the local data archive context. REGARDS configuration and ingestion of Solar Orbiter data and other MEDOC datasets (see Fig. 2) are still ongoing (SDO data have hundreds of millions of records). As for SiTools2, MEDOC has developed a Python client for the REGARDS API, and REGARDS also provides an OpenSearch interface.

Some MEDOC-specific derived data products are also provided through a Virtual Observatory (VO) API with the EuroPlanet Table Access Protocol (EPN-TAP), thanks to a DaCHS server. These products have then been made available through several Virtual Observatories, such as VESPA (Virtual European Solar and Planetary Access, Erard et al. 2018, EuroPlanet 2024 Research Infrastructure) and Solar-VO (SOLARNET H2020 project).

Four products are provided to the ESA Space Weather Portal[‡], including web pages with latest data and query forms to the archive, and a custom API. By ESA request, the web page and API had to be protected by the ESA single-sign-on system, but registration is open to everyone. All datasets have a Space Physics Archive Search and Extract (SPASE, Roberts et al. 2018) description file, which has been submitted to ESA.

Finally, MEDOC hosts data output from some numerical simulations runs: solar wind profiles (hydrodynamic models from LPP and LESIA laboratories), a solar eruption (3D MHD, also from LPP and LESIA), and the Solar wind with cosmic rays scattering (3D MHD, from AIM laboratory). The custom web pages hosting these results are planned to be migrated to a new Galactica [§] node; Galactica is built following the International Virtual Observatory (IVOA) Theory Data Model.

[‡] `https://swe.ssa.esa.int/`
[§] `http://www.galactica-simulations.eu/db/`

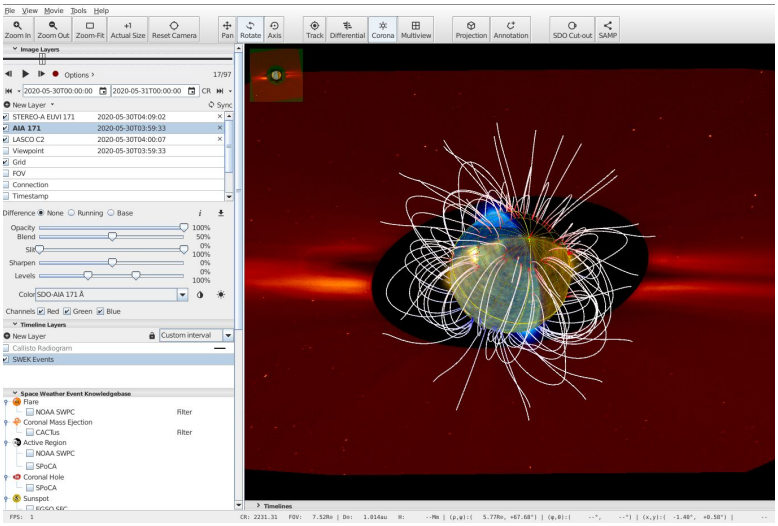

Fig. 3. HelioViewer Java client (JHelioViewer) with data from the MEDOC server, including coronal observations from several viewpoints, coronagraphic data, and magnetic field lines from a Potential Field Source Surface (PFSS) extrapolation.

4 Data exploration and visualization

The large data volumes available in heliophysics, with multi-scale data and an important time variability, make it very useful to have dedicated tools for multi-instrument data exploration and visualization. This is why ESA and NASA have jointly developed the HelioViewer system, including a JPEG 2000 Interactive Protocol (JPIP) server, and web and Java clients (among other clients developed by the scientific community); see Fig. 3 and Müller et al. (2017). MEDOC hosts a server instance, with a full mirror of the data (127 TB). The Java client has a Simple Application Messaging Protocol (SAMP) interface, allowing for example users to get FITS data directly from the ESA Solar Orbiter archive, and to combine them with other data in HelioViewer. A HAPI (Heliophysics data API, Weigel et al. 2021) functionality is in development, for the servers and clients, to provide a better way of exchanging information about data availability.

5 Conclusion

MEDOC provides open data and tools for the solar physics, heliophysics, and space weather communities. Data are using standard file formats (mainly FITS), and efforts are made to provide APIs and complete and useful metadata. Most datasets also have Digital Object Identifiers (DOIs), for their persistent identification. Efforts are made to adopt standards (in particular the IVOA ones), although much work is still to be done in the case of tools that are not fully compliant with these standards yet.

References

Erard, S., Cecconi, B., Le Sidaner, P., et al. 2018, Planet. Space Sci., 150, 65

Müller, D., Nicula, B., Felix, S., et al. 2017, A&A, 606, A10

Roberts, D. A., Thieman, J., Génot, V., et al. 2018, Space Weather, 16, 1899

Weigel, R. S., Vandegriff, J., Faden, J., et al. 2021, Journal of Geophysical Research (Space Physics), 126, e29534