Open solar data, data products, and tools at MEDOC

É. Buchlin, S. Caminade, S. Parenti, B. Perri, F. Auchère,
N. Traoré, K. Ashkar, D. Leung

IHDEA 2024, Madrid, 2024-10-17

What is MEDOC ? https://idoc.osups.universite-paris-saclay.fr/MEDOC/

A *data and operations center* for *solar physics space missions* :

- Created in 1996, as *European SoHO data and operations center*
- Since then, *many other solar data sets* (STEREO, SDO...), but also *derived data products* and *tools* to use data → facilitate data exploitation
- SoHO/GOLF and Solar Orbiter/SPICE *operations*

- Funding from CNES, CNRS, and Université Paris-Saclay
- Projects/collaborations with ESA, EC, CNES, CDPP...

What is MEDOC ? https://idoc.osups.universite-paris-saclay.fr/MEDOC/

A *data and operations center* for *solar physics space missions* :

- Created in 1996, as *European SoHO data and operations center*
- Since then, *many other solar data sets* (STEREO, SDO...), but also *derived data products* and *tools* to use data → facilitate data exploitation
- SoHO/GOLF and Solar Orbiter/SPICE *operations*

- Funding from CNES, CNRS, and Université Paris-Saclay
- Projects/collaborations with ESA, EC, CNES, CDPP...

Officially CNRS-recognized "observation services" for heliophysics data

- MEDOC : space solar/heliospheric data (remote-sensing observations)
- CDPP : solar system natural plasmas (in-situ measurements)
- *3SOLEIL* : ground solar data (VL, radio, cosmic rays)
- APIS (planetary auroral observations), MASER (radio), CLIMSO, STORMS, ISGI.

...offering :

- Access to *data* and products derived from observation data
- *Tools and services* to access and exploit data

...in an effort to make data and metadata

Officially CNRS-recognized "observation services" for heliophysics data

- MEDOC : space solar/heliospheric data (remote-sensing observations)
- CDPP : solar system natural plasmas (in-situ measurements)
- *3SOLEIL* : ground solar data (VL, radio, cosmic rays)
- APIS (planetary auroral observations), MASER (radio), CLIMSO, STORMS, ISGI.

...offering :

- Access to *data* and products derived from observation data
- *Tools and services* to access and exploit data

...in an effort to make data and metadata

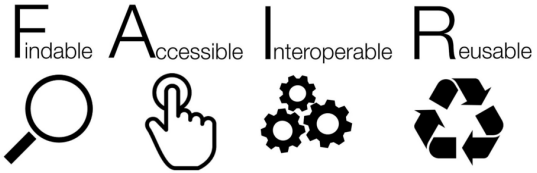
Officially CNRS-recognized "observation services" for heliophysics data

- MEDOC : space solar/heliospheric data (remote-sensing observations)
- CDPP : solar system natural plasmas (in-situ measurements)
- *3SOLEIL* : ground solar data (VL, radio, cosmic rays)
- APIS (planetary auroral observations), MASER (radio), CLIMSO, STORMS, ISGI.

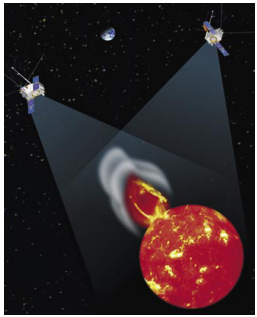
...offering :

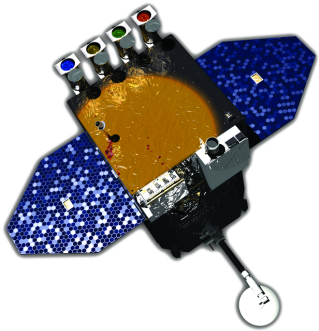
- Access to *data* and products derived from observation data
- *Tools and services* to access and exploit data

...in an effort to make data and metadata



What does MEDOC provide ? Observation data

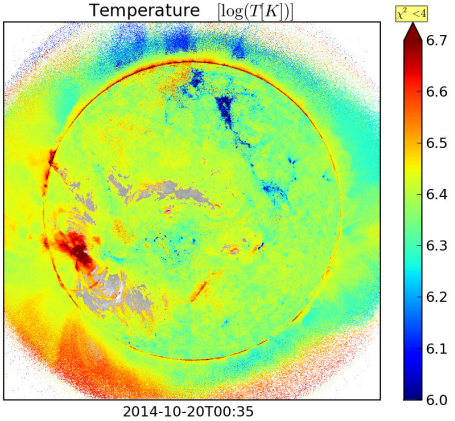




- *Current missions* : SoHO (1995–), STEREO/SECCHI (2006–), SDO (2010–) : $>$ 700 TB
  Coming soon : Solar Orbiter (2020–)
- PICARD (2010–2014 ; *CNES main archive*), TRACE (1998–2010)...

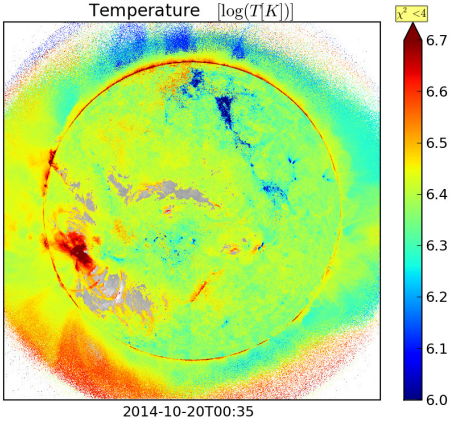
Derived data products (computed from observation data)

- Thermal structure of the corona (DEM parameters)
- Synchronous synoptic maps of the corona and photosphere
- Electric currents in active regions
- Movies, spectral atlases, catalogs...



Derived data products (computed from observation data)

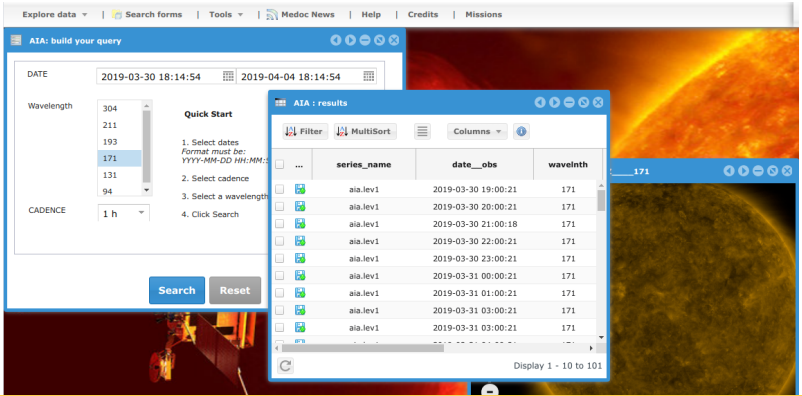
- Thermal structure of the corona (DEM parameters)
- Synchronous synoptic maps of the corona and photosphere
- Electric currents in active regions
- Movies, spectral atlases, catalogs...



Some of these are computed *several times per day*, just a *few hours after the observation*.

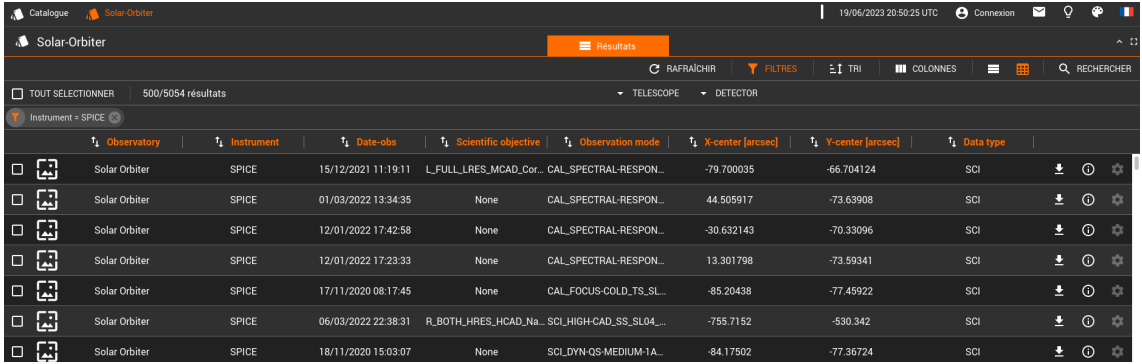
Main interface : SiTools2

- Developed (2010–2017) by AKKA for CNES, Java, ExtJS
- REST API, for which we have developed IDL and Python clients
- Web interface
- Not maintained anymore by CNES → migration needed.



New interface : REGARDS

- Developed by CS for CNES
- Built on Spring microservices, implements OAIS recommendation
- REST API (for which we have developed a Python client)
- Web interface (React+Redux)



How we are migrating to REGARDS

- *Setting up all datasets* : displayed columns, query forms, filters....
  Now using a single data model.
- *Ingesting* data :
  - Build OAIS SIP packets from pre-existing databases (SOHO, STEREO, SDO...) or from TAP interface (ESA Solar Orbiter archive).
  - Submit them to the ingestion queue using AMQP (RabbitMQ)
  - After ingestion, the products are harvested into ElasticSearch
- REGARDS can in principle ingest data using OpenSearch, but is this supported by another heliophysics archive ?

Some challenges :

- A lot of "hidden" work before datasets are migrated or ingested.
- Currently no good technical solution for ingesting our > 300M SDO records
- REGARDS, as a generic CNES tool, doesn't support heliophysics API standards out-of-the-box.

How we are migrating to REGARDS

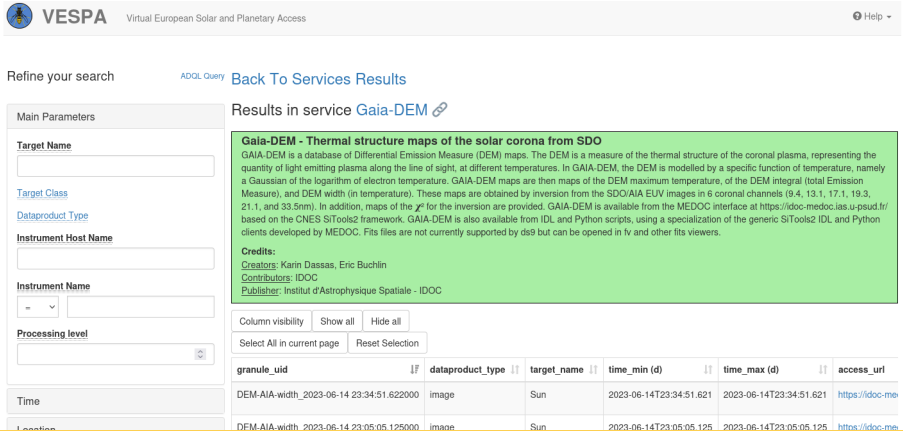
- *Setting up all datasets* : displayed columns, query forms, filters....
  Now using a single data model.
- *Ingesting* data :
  - Build OAIS SIP packets from pre-existing databases (SOHO, STEREO, SDO...) or from TAP interface (ESA Solar Orbiter archive).
  - Submit them to the ingestion queue using AMQP (RabbitMQ)
  - After ingestion, the products are harvested into ElasticSearch
- REGARDS can in principle ingest data using OpenSearch, but is this supported by another heliophysics archive ?

Some challenges :

- A lot of "hidden" work before datasets are migrated or ingested.
- Currently no good technical solution for ingesting our $>$ 300M SDO records
- REGARDS, as a generic CNES tool, doesn't support heliophysics API standards out-of-the-box.

EPN-TAP VO services

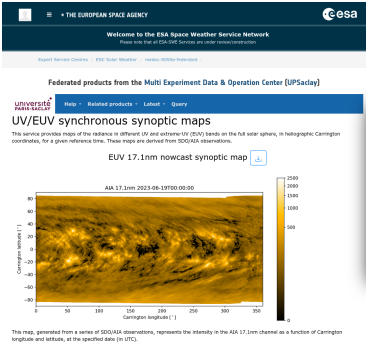
- A few MEDOC-specific datasets are available through EPN-TAP using a DaCHS server
- Then they are available from *Virtual Observatories* : VESPA and Solar-VO
- Maybe soon in VSO, or waiting for VSO 2.0 ?

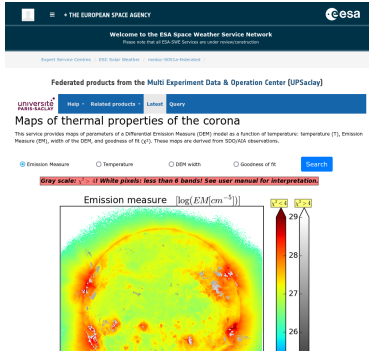


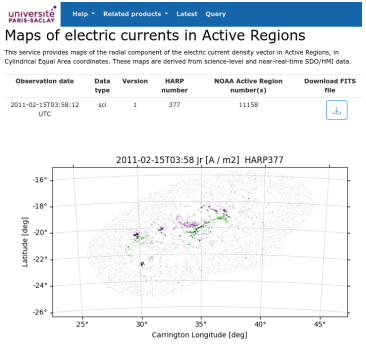
Products for the ESA Space Weather Portal https://swe.ssa.esa.int/

MEDOC provides 4 products (maps derived from data) to ESA's Space Weather network portal (ESA contract numbers 4000128012/19/D/MRP and 4000134036/21/D/MRP) :

- Web apps with well-determined functionalities and layout
- API (freely specified — HAPI was not ready when we started)
- All datasets have a SPASE description file submitted to ESA (not in any registry yet ?)





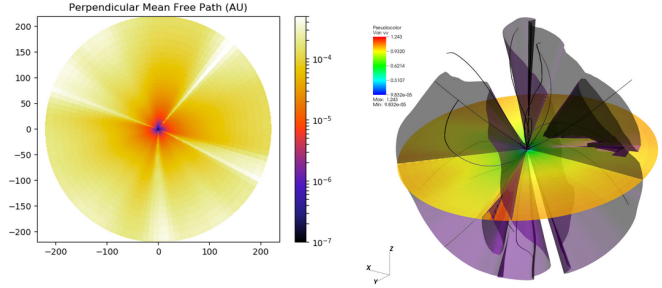


Data from numerical simulation runs

- Solar wind profiles (VP code, 1D hydrodynamics)
- Solar eruption (OHM code, 3D MHD)
- Solar wind with cosmic rays scattering (PLUTO code, 3D MHD).

Will soon be provided through the new MEDOC node of Galactica (CEA/AIM, OP/VO-PDC, Heidelberg)

Galactica is built following IVOA Theory Data Model, and uses Django, Celery, RabbitMQ

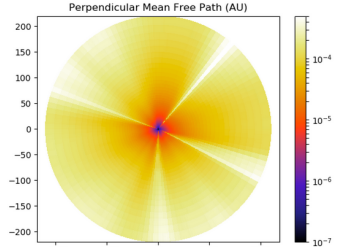


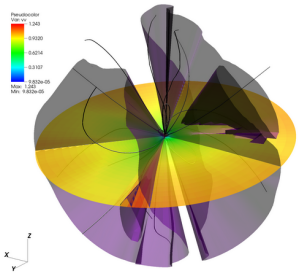
Data from numerical simulation runs

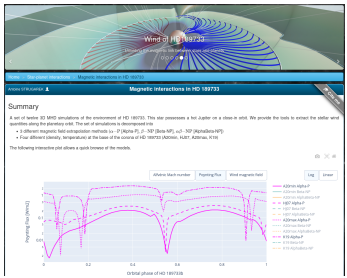
- Solar wind profiles (VP code, 1D hydrodynamics)
- Solar eruption (OHM code, 3D MHD)
- Solar wind with cosmic rays scattering (PLUTO code, 3D MHD).

Will soon be provided through the new MEDOC node of Galactica (CEA/AIM, OP/VO-PDC, Heidelberg)

Galactica is built following IVOA Theory Data Model, and uses Django, Celery, RabbitMQ

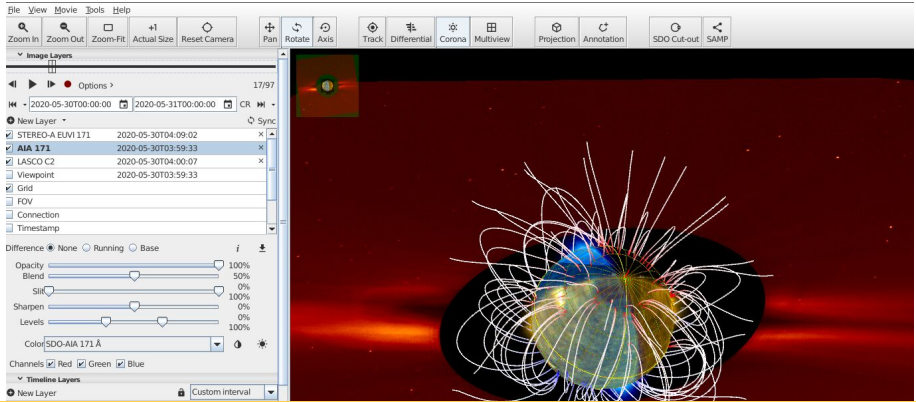






Data exploration / visualization : HelioViewer (ESA and NASA)

- HelioViewer JPIP (JPEG2000) server at MEDOC. New (ESA, ROB) : HAPI interface.
- Full mirror of NASA-GSFC HelioViewer data (100TB data).
- JHelioViewer SAMP interface, can be used e.g. to get FITS data from ESA Solar Orbiter archive



Conclusion

- MEDOC provides data and tools for the community
  - All are open, available with no registration
- Efforts to provide APIs, complete metadata, DOIs...
  as well as to adopt standards (in particular IVOA ones)
- But many tools are not fully compliant with these standards... Still much work to do

Conclusion

- MEDOC provides data and tools for the community
  - All are open, available with no registration
- Efforts to provide APIs, complete metadata, DOIs...
  as well as to adopt standards (in particular IVOA ones)
- But many tools are not fully compliant with these standards... Still much work to do.