

Federation of Solar Data (FSD): A Framework for FAIR and Interoperable Solar Physics Data

A. Wiśniewska¹
Collaborators

¹Astronomical Institute of Slovak Academy of Sciences

Abstract

The Federation of Solar Data (FSD) project establishes a unified framework for archiving, curation, and dissemination of solar observational and numerical simulation data following FAIR (Findable, Accessible, Interoperable, Reusable) principles. By implementing SOLARNET metadata recommendations and International Virtual Observatory Alliance (IVOA) standards, heterogeneous datasets from multiple solar instruments and radiative magnetohydrodynamic simulations are integrated into Virtual Observatory infrastructures. The project enables advanced data discovery, cross-instrument analysis, and prepares data workflows for future large-scale facilities such as the European Solar Telescope.

1 Introduction

Modern solar physics produces rapidly increasing volumes of observational and simulation data. These datasets are often distributed across multiple institutions, stored in heterogeneous formats, and described using non-uniform metadata structures.

The Federation of Solar Data (FSD) project addresses these challenges by federating solar datasets into interoperable infrastructures compliant with Virtual Observatory standards. The project focuses on harmonization of metadata, long-term preservation, and open dissemination of solar data products across the astrophysical community.

2 Project Objectives

The main objectives of the FSD project are:

- Implementation of SOLARNET metadata standards.
- Adoption of IVOA and EPN-TAP interoperability protocols.

- Dissemination of solar datasets via the Solarnet Virtual Observatory (SVO) and/or IVOA.
- Integration of observational and simulation datasets.
- Alignment with FAIR and EOSC principles.

3 Data Sources

FSD integrates both observational and numerical datasets.

3.1 Observational Data

Data originate from multiple ground-based solar facilities including:

- Swedish Solar Telescope (CRISP, CHROMIS)
- Dunn Solar Telescope (IBIS)
- GREGOR telescope (GRIS)
- ZIMPOL spectropolarimetric systems
- CoMP-S instrument

3.2 Numerical Simulations

Radiative magnetohydrodynamic simulations included in the federation:

- CO5BOLD
- Bifrost
- MURaM
- Pencil Code
- MANCHA3D
- DISPATCH framework

4 Methodology

Datasets are converted into SOLARNET-compliant formats and exposed through Table Access Protocol (TAP) services. Metadata harmonization enables automated discovery and programmatic access using VO-compatible clients.

Implementation of EPN-TAP ensures standardized descriptions of temporal, spectral, and spatial coverage, allowing cross-dataset scientific searches.

5 Scientific Impact and Deliverables

The scientific impact of FSD derives from transforming previously isolated datasets into interoperable scientific resources.

Key project deliverables include:

- FAIR-compliant solar datasets accessible via VO services.
- Harmonized metadata standards across instruments.
- Integration of solar data into SVO and astrophysical archives.
- Dissemination of r(M)HD simulations to the wider community.
- Preparation of workflows applicable to EST-era data volumes.
- Community training and workshops supporting VO adoption.

These developments enable reproducible research, multi-instrument analysis, and long-term preservation of high-value solar observations.

6 Instrument Contributions

This section is intentionally modular. Each partner may extend their subsection independently without modifying the global manuscript structure.

6.1 CRISP and CHROMIS — Swedish Solar Telescope

Describe dataset preparation, pipeline updates, metadata implementation, and integration into SVO.

6.2 IBIS and IBIS 2.0 — Dunn Solar Telescope

Describe archive expansion, metadata standardization, FAIR compliance, and dissemination workflows.

6.3 ZIMPOL Spectropolarimetry

Describe incorporation of high-precision spectropolarimetric datasets, database upgrades, and VO accessibility.

6.4 GRIS — GREGOR Telescope

Describe ingestion of upgraded GRIS datasets, format conversion, and interoperability implementation.

6.5 CoMP-S Instrument

6.5.1 AISAS: task 1

Task title: Preparation of the CoMP-S instrument data to be completely compatible with the SOLARNET Metadata Recommendations for Solar Observations (Version May 13, 2025 - SOLARNET=1.0).

The CoMP-S instrument data of level-1 were analyzed in relation to the SOLARNET Metadata Recommendations for Solar Observations (Version May 13, 2025) Then the list of the inevitable changes in order to prepare those data to be compatible with the actual standard specified in this document was prepared. Later on, step by step, inevitable changes of the CoMP-S instrument data of level-1 keywords in their FITS files, located in the primary and secondary headers, were performed. It should be noted here that these changes were quite large completely changing the policy how to separate the keywords between the primary and secondary headers. All this was needed to be in line with the SOLARNET metadata standard. A completely new part of information was included regarding the WCS format of orientation of the CoMP-S coronagraphic images which is now fully compatible with the developed solar WCS standard for the FITS files. Additionally, the quick-look previews of the data were prepared showing the complete sets of 4 polarimetric measurements for each wavelength step to which the Lyot filter of the instrument was tuned.

The headers of the FITS files were searched for the required information to be selected for the transfer table to fill the EPNcore table with the inevitably required and also recommended keywords from the CoMP-S instrument data - level-1 FITS files. This work was in fact running in both directions as some keywords were added to the CoMP-S instrument data

- level-1 FITS files based on ideas received from our EPNcore table template and such tables prepared by colleagues for other solar instruments.

For this task 1 of AISAS in the OSCARS project the CoMP-S instrument data - level-1 FITS files were selected.

The CoMP-S instrument data - level-1 data cover real observations of the solar prominences and emission corona performed at the Lomnický Štít Observatory (LSO) using the ZEISS 200/3000/4000 coronagraph to which the CoMP-S is attached. This level-1 is the level containing the raw data of observations prepared in format of the FITS files where spectropolarimetric data are stored in the binary extensions to the basics FITS format standard. Moreover, the FITS files major and minor headers contain the complete auxiliary information needed to describe the data stored.

The main results of this task is that the CoMP-S instrument level-1 data are completely in line with the specific standards of the FITS basic format, its binary extension description, the WCS standard for the solar observations and the SOLARNET Metadata Recommendations for Solar Observations (Version May 13, 2025).

The main deliverable of the task is that it prepared the CoMP-S level-1 data in the format fully compatible with its posting in the SOLARNET archive.

This deliverable of this task allowed that work of the complementary AISAS task - task 2 - was able to perfume the actual posting of the datasets of CoMP-S instrument in the SOLARNET Virtual Observatory archive.

6.5.2 AISAS: task 2

Task title: Making the CoMP-S data archive accessible through SVO and the SOLARNET archive.

Hereby, the list of actions and work carried out in frame of this task is given:

- Installation of DACHS software and Ubuntu upgrade on a dedicated data storage server
- Configuration of a PostgreSQL database cluster based on gavo DB requirements
- verification with GAVO DB support whether it is possible to import FITS files with multiple extensions - in reply fix was received
- modification of the Resource Descriptor (RD) template file for solar data, for FITS files from LSO/CoMP-S instruments with multiple binary extensions

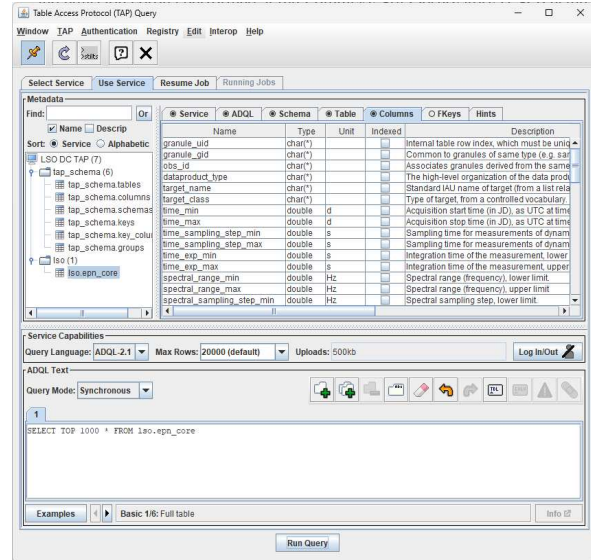


Figure 1: TAP query for a table lso.epn-core from TOPCAT tool.

	granule_uid	obs_id	dataproduct_type	target	
1	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
2	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
3	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
4	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
5	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
6	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
7	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
8	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
9	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
10	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
11	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
12	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
13	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
14	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
15	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
16	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
17	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
18	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
19	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run
20	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	lso_comp_s_level1_20250912_253817_waw0654	rs	Run

Figure 2: TAP query results for table lso.epn-core from TOPCAT tool.

- testing data import using real LSO/CoMP-S Level-1 FITS files with binary extensions. The issue with importing the DATE-BEG and DATE-END keywords has been resolved
- installation and configuration of a local web server for displaying thumbnails via web links from the LSO.EPN-CORE database table
- configuration of the web server for downloading individual FITS files via web links from the LSO.EPN-CORE database table
- connecting the disk storage to DB server to access the data for import
- connecting disk storage to the database server to access data intended for import
- testing the import of compressed FITS file with .fz suffix into GAVO DB
- Installation of the TOPCAT tool and testing of local data publication to ensure their visibility via TOPCAT
- validating our lso.dc authority before adding it to the Registry of Registries
- verification of our lso.dc authority before its addition to the Register of Registers
- elimination of deficiencies identified during the registration of our registry in Registry of Registries RofR
- verification of the availability of our imported data in the database via the TAP protocol using the TOPCAT tool
- provision of information to the SVO administrator required for importing our data into the SVO database
- initial data import into the SVO DB, data verification, data searching and display of previews. Proposals for changes to the SVO DB sent to the database administrator.
- modifying and improving data import scripts together with the database administrator.
- final import of LSO CoMP-S level 1 data into the SVO database

For this task 2 of AISAS in the OSCARS project the CoMP-S instrument data - level-1 FITS files were selected.

The CoMP-S instrument data - level-1 data cover real observations of the solar prominences and emission corona performed at the Lomnický štít Observatory (LSO) using the ZEISS 200/3000/4000 corona-graph to which the CoMP-S is attached. This level-1 is the level containing the raw data of observations prepared in format of the FITS files where spectropolarimetric data are stored in the binary extensions to the basics FITS format standard. Moreover, the FITS files major and minor headers contain the complete auxiliary information needed to describe the data stored.

Software used for this purpose:

- ubuntu 24.04 LTS
- DACHS server "gavodachs2-server" -> version 2.9.1+dfsg-1, upgraded to version 2.11
- other software packages required by DACHS server:
 - gavodachs2-server - version 2.9.1+dfsg-1,
 - PostgreSQL DB upgrade from 14 to 16.6,
 - postgresql-16-pgsphere,
 - postgresql-16-q3c,
 - queryparser - tool for parsing and processing of PostgreSQL (from python) and translation of ADQL SELECT-like queries,
 - pg-sphere - pgSphere is a PostgreSQL extension for spherical geometry, and
 - TOPCAT (Tool for Operations on Catalogues And Tables) tool was used to search database datasets via the TAP protocol.

The main result of this task is that the Level 1 data from the CoMP-S instrument comply with the EPN-Core metadata dictionary specification and are accessible via the EPN TAP protocol, as well as through SVO and the SOLARNET archive.

The main deliverable of the task is that it prepared the CoMP-S level-1 data in the format fully compatible with its posting in the SOLARNET archive.

This deliverable of this task allowed that work of the complementary AISAS task 2 was able to perform the actual posting of the datasets of CoMP-S instrument in the SOLARNET Virtual Observatory archive.

6.6 Numerical Simulation Contributions

Partners contributing simulations should describe:

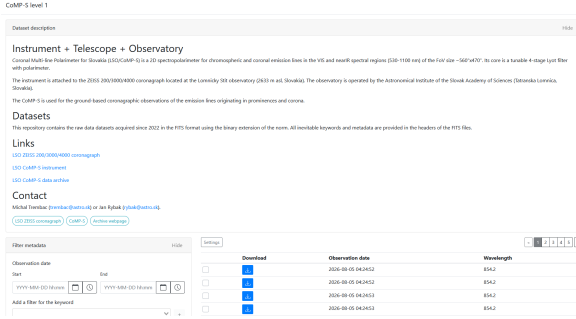


Figure 3: LSO/CoMP-S Level-1 dataset displayed via SVO web client.

- simulation code,
- physical setup,
- data products,
- metadata conversion,
- VO dissemination strategy.

7 Digital Infrastructure

FSD integrates distributed archives, metadata services, and storage solutions aligned with the European Open Science Cloud (EOSC). Authentication and authorization infrastructures enable controlled yet open scientific access.

The developed tools provide a prototype for future solar data lakes and large-scale data federation initiatives.

8 Conclusions

The Federation of Solar Data project establishes a scalable framework for solar data federation based on FAIR principles and Virtual Observatory standards. The resulting infrastructure enhances scientific reuse, interoperability, and preparedness for next-generation solar facilities.

Acknowledgements

This work is supported by the OSCARS Open Call programme.

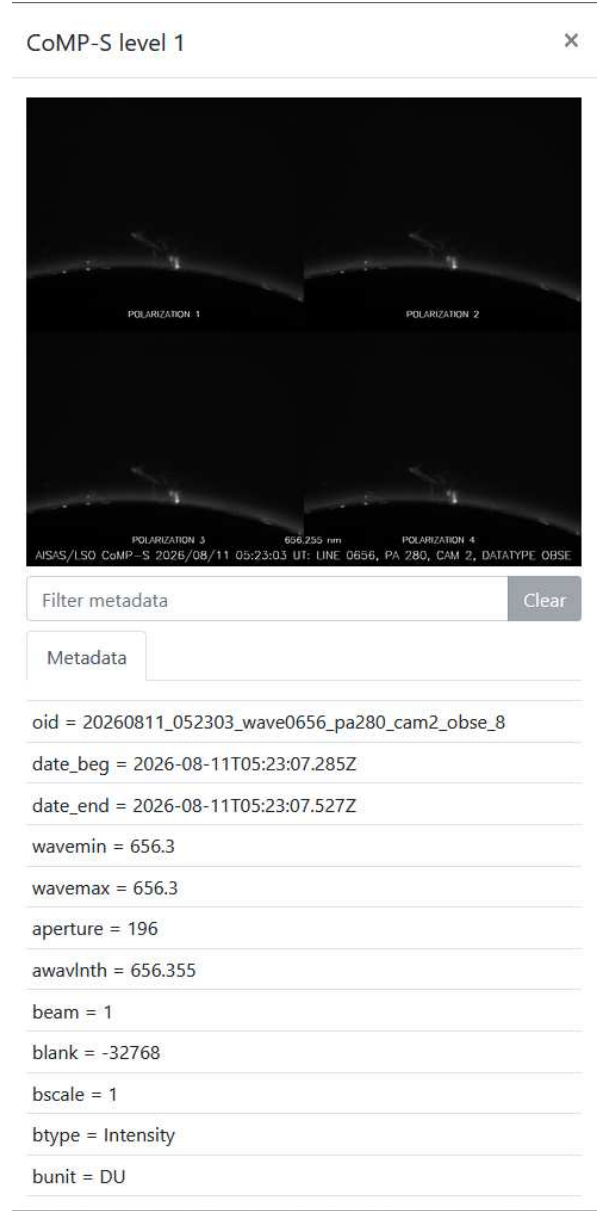


Figure 4: LSO/CoMP-S level-1 data preview displayed via SVO web client.

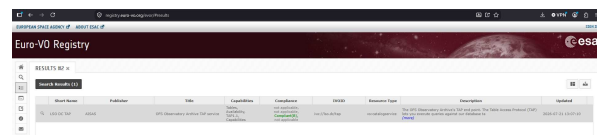


Figure 5: LSO/CoMP-S Iso.dc authority registered in Euro-VO Registry together with GAVO Registry.