

Implementation of the Data Management Plan (DMP)



Goal is to provide FAIR based data for EUROfusion (related to Grant deliverable).
Charge for 2025 is to

- **Provide a searchable catalogue/database** of metadata (waveforms) from the participating sites. (Scenario A)
- **Demonstrate direct data access** of a subset of experimental data for user applications to run on. (Prototyping Scenario B).
- **Investigate** (and pending available resources) develop the technology to integrate **modelling/simulation data** as a "facility" of its own

Activity is divided into **Core services** (PSNC) providing infrastructure and **Sites** (**AUG, COMPASS-U, JET, MAST-U, TCV and WEST**) providing data mappings and remote data access.

With the long term data storage facility (LTDSF) and the ability to mint PID's we are have the tools to support the longer term vision towards a one stop facility for researching, accessing, processing, analysing, and sharing experimental and modelling data.

Implementation of the Data Management Plan (DMP)



Infrastructure is in place!

- A UDA (UKAEA) based client/server installation is available and tested on all sites.
- Data ingestion and curation procedures and protocols have been tested for Scenario A data (metadata services) for all devices
- UDA has been updated and is being tested with new security enhancements to allow for Authentication and Authorisation (needed for Scenario B data releases)
- Performance issues with large data volumes has been resolved
- Future work is related to adaptation to user needs and further performance improvements.

Metadata - Waveforms, etc (Scenario A) is ready for production services

- Ready for production services: Will be launched on the new gateway (faster access to HW is being explored).
- EUROfusion users will initial have access to: AUG: (11,500 discharges), WEST (~1,200), TCV: (Near full dataset), JET (TBD: on request mapping tool), MAST-/U & Compass-/U (pending release agreement)
 - Some embargos and data restrictions may apply

Direct data access has been demonstrated for a number of use cases (scenario B):

- Remote access tools are available using UDA/IMAS
- A set of user needs representing different TSVVs needs have been defined and tested.
- Relevant data mappings are being developed and have been demonstrated for select devices.

Initial testing of integration with modelling/simulation data has started.

- Strategy is to use SimDB and integrate with UDA access
- Allow users to log/share simulation results through the catalogue
- The availability of the long term data storage facility (LTDSF) is of key importance to this



2025 activities: release and expand

- Future adaptation to user needs and further performance improvements of the infrastructure elements
- Continuous updates to the released metadata catalogue
- Extended set of use cases supported for direct data access on a full(er) set of devices
- Demonstrate integration of IMAS based modelling /simulation data into the existing infrastructure.

Collaborations with WP PrIO: (2026-27 continuation? DMP new home?)

- Common postprocessing of diagnostic data - provide core_profiles using IMAS based IDA toolset
- EUROfusion databases: uses the catalogue backend for data releases
- (pedestal, disruption,..)

Continued collaborations with ITER (and UKAEA) on UDA, SimDB and IMAS. Should be on CCFE priority list

Implementation of the DMP (scenarios)



The data management plan defines 4 scenarios/stages of increasing ambition

- Scenario A: making metadata only available and searchable using IMAS data subsets for interoperable definitions of quantities [F,(I)]. **Production ready**
- Scenario B: adds to Scenario A by allowing a subset of the data to be accessed using common tools (UDA). Facilities are responsible for the access level and qualification of data through the data mappings [F,A,I,(R)]. **Demonstrated – being expanded**
- Scenario C: builds on the previous stages and allows for enhanced data provenance and referencing through PID's [F,A,I,R]. **Defer. However, LTDSF and PID minting makes it a very interesting 2026-27 prospect.**
- Scenario D: adds a lightweight layer for open access to non-embargoed metadata and where allowed by the facilities also data access for export in human readable formats (CSV files) [F,A,I,R] and open. **Defer.**

FAIR - findability, accessibility, interoperability, and reusability