2026-27 Goals Data management plan

**Grant deliverable 2027**

* **Fully implement the EUROfusion Data Management Plan** by Q4 2027, ensuring compliance with FAIR principles across all IMAS enabled EUROfusion devices and sponsored modeling activities.

**Deliverables:**

* **Develop and deploy an IMAS integrated data-sharing environment** by Q4 2026, standardizing experimental and simulation data in IMAS formats to enable seamless interoperability using remote data access and local databases.
* **Establishing open data access** for select long-term storage datasets by Q4 2026, ensuring secure, structured availability for research and analysis.

Extending scope to cover AI/ML data and model networks.

* **Deliver AI/ML-ready tools and datasets** based on IMAS ontologies by Q4 2026, supporting use cases such as surrogate model development and experimental data synthesis, with validated applications supporting each of the relevant ongoing AI/ML based projects.

Indicative resource needs Core services (2026: 24PM, 2027:24PM), Site services (9PM, 9PM) per site [AUG, COMPASS, TCV, WEST], with collaborative agreement [MAST/MAST-U, and JET?] + management. Questionmarks: JT60-SA (IMAS integration status?), W-7X (IMASIFICATION status?). Integration with AI/ML needs additional resources but depends on scope and structure.

|  |  |  |
| --- | --- | --- |
| Activity | 2026 (PM) | 2027 (PM) |
| Site Services (AUG, COMPASS, TCV, WEST) | 36 | 36 |
| Site Services (MAST/MAST-U cooperation) | (0) | (0) |
| Site Site Services(JET) | (9) | (9) |
| Core Services | 24 | 24 |
| Management | 3 | 3 |
| Totals | 63 (72) | 63 (72) |

Possible amendments (each on the same level of investment as a single site)

* W-7X could join towards some scenario levels depending on IMASifcation status.
* JT60-SA could be integrated - EUROfusion need to have a clear policy on JT-60SA IMASification
* AI/ML - potential use of the DMP infrastructure to deploy and distribute related data and models. Actual data curation and regeneration of models in relevant tasks or projects.

**Implementation**

Use the IMAS toolset for sharing data that are either from experimental devices or born from simulations (modelling experimental discharges or born digital) we will promote scenario C for this assuming a LTDSF is available and that the ability to mint PID’s has been installed. Extending Scenario D is straightforward, assuming the LTDSF can promote a publicly available partition.

Initial support for AI/ML activities needs to be done in close collaboration with users and a requirements/needs capture or prototyping should be performed with the AI/ML pilots under 2025. Distribution or deployment of AI/ML data and models could be promoted through the dashboard interfaces. Requirements capture and prototyping would be needed in 2025. Potentially NEW resources for 2026-27.

**Core Elements: (24 PM 2026, 24 PM 2027)**

* Development, Maintenance and support of Dashboard/ catalogue tools and UI. Including a potential move to document-based databases (currently only prototyped)
* Development, Maintenance and support for Remote access tools for ingestion and data access (UDA based)
* Further development of batch processing functionality (e.g. CLI access to resources
* Integration with local IMAS data storage (including SimDB)
* integration with EUROfusion AAI - development of quasi AAI
* PID minting
* Code refractoring/migration towards Python based tool set to align with general fusion developments including ITER

Core workflow amendments (3pm)

* IDA mappings support for core\_profiles, (ongoing with deliverable 2025 – support needed 2026-27)
* Core\_sources, Distributions (ASCOT/RABBIT etc)

Site specific (CEA 9PM 2026, 2027) Suggest 9PM per site to cover some additional development needs although it is a bit current resource needs.

* Maintenance and support for local remote access server (UDA)
* Local data mappings to support the fulls et of use cases
* Development, support and maintenance for local ingestion workflow
* Improved provenance capture

Management (3PM)

General concerns:

* Access policies and methodologies agreed (surprisingly slow)
* Quality and resolution management of delivered data

Risks

* Long term UDA support
* Long term SimDB support
* Site support for data mappings i(n particular JET)

Range of IDS requested:

* Summary, dataset\_description
* Magnetics, pf\_active, pf\_passive, tf and wall + [iron\_core], constraints (MSE etc)
* Equilibrium, bolometer, Thomson\_scattering, soft\_xrays, ec\_launchers, ece, ic\_antennas, nbi
* core\_profiles, core\_sources, [core\_transport], distributions