XS-ABILITY

Accessing hard-to-reach areas with Advanced and Breakthrough Innovation for reLiable In-situ characterization of a facility

Deliverable 1.2
Data Management Plan

Document Identification & Project Details			
Project Acronym	XS-ABILITY	***Grant agreement***	101166392
Project start date	01/10/2024	***Project duration***	36 months
Document status	Draft	***Related WP***	WP1
Version	V.1	***Reference***	D1.2
Call identifier	HORIZON-EURATOM-2023-NRT-01-07		

Funded by the European Union

This project has received funding from the European Union (HORIZON-EURATOM-2023-NRT-01), under grant agreement n° 101166392. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the granting authority. Neither the European Union nor the granting authority can be held responsible for them.

HISTORY & DETAILS

Deliverable details	
Due date	31/12/24
Actual submission date	31/01/25
Lead beneficiary	CEA
Author(s)	Maugan MICHEL
Contributor(s)	N/A
Dissemination level	Public

Author	Version	Date	Reason for change	Release
Maugan MICHEL	1	31/01/25	N/A	N/A

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

Acronym	Description
AB	Advisory Board
AP	Associated Partner
CA	Consortium Agreement
DMP	Data Management Plan
DoA	Description of the Action
EB	Executive Board
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable and Re-usable
GA	Grant Agreement
PC	Project Coordinator
PM	Project Manager
PR	Progress Report
TK	Task
WP	Work Package
WPL	Work Package Leader

EXECUTIVE SUMMARY

The XS-ABILITY Data Management Plan (DMP) follows the Horizon Europe DMP template that was designed to be applied to any European project that produces, collects or processes research data.

This DMP deliverable describes the data management principles and strategies, tools and XS-ABILITY data that will be produced as part of the project activities and that are relevant to be included in the DMP. The consortium will also aim at open access when publishing papers and articles.

The DMP is a living document to be updated as the implementation of the project progresses and when significant changes occur.

KEY WORDS

Data Management Plan; FAIR, Findable, Accessible, Interoperable, Re-usable; data

1. INTRODUCTION

1.1. Why would I want to read this deliverable?

It provides an easy overview of research data the project is expected to generate, the types and formats of this data, and how this data is processed and stored to make them findable, accessible, interoperable and reusable, according to the principles of FAIR data management. The purpose of the DMP is to contribute to good data handling during the project's lifetime, and to describe how such data will be curated and preserved.

1.2. Intended readership/users?

Internally to the project:

- All project participants who are responsible for, or in any way involved with, data collection and data handling can use this document for instructions on how to handle, store and process data.
- All project participants can use this document to get an overview of all data collected in the project and how this is processed and stored.

External audience:

- All relevant stakeholders who are interested in XS-ABILITY related activities and research topics can use this document to get an overview of the data collected in the project, how to access this data, and, if applicable, how to re-use these data in their own activities

1.3. Objectives and scope of the document

The DMP describes the data management life cycle for the data to be collected, processed and/or generated by the XS-ABILITY project, as a European project. The DMP aims at defining the management strategy of data generated during the project with the purpose to making research data Findable, Accessible, Interoperable and Re-usable (FAIR).

1.4. Structure of the deliverable

The document is structured following the guideline of Horizon Europe programme on FAIR Data Management in Horizon Europe including the following information:

- DMP guiding principles and strategy
- Description of XS-ABILITY type of data
- Description of FAIR data characteristics including DMP Review Process & data inventory
- Allocation of resources
- Data Security
- Ethical Aspects
- Conclusions

2. DATA SUMMARY

The XS-ABILITY DMP aims to provide a strategy for managing key data generated and collected during the project and optimize access to and re-use of research data. The DMP is intended to be a 'living' document that will outline how the XS-ABILITY research data will be handled during and after the project, and so it will be reviewed and updated at regular intervals.

All European Union funded projects must try to disseminate as much information as possible. In this regard, the main purpose of the DMP is to ensure the accessibility and intelligibility of the data generated during the XS-ABILITY project in order to disseminate information the best we can. Each dataset created during the project will be assessed and categorized as open, embargo or restricted by the owners of the content of the dataset.

All the datasets, regardless of their categorization, will be stored in each of the participant entities databases. In addition, those categorized as open or embargo will be publicly shared (in the case of embargo, after the embargo period is over) through the public section of the project website and ZENODO (https://zenodo.org/).

ZENODO is an open access repository for all fields of science that allows uploading any kind of data file formats, which is recommended by the Open Access Infrastructure for Research in Europe (OpenAIRE).

2.1. DMP guiding principle

The DMP of XS-ABILITY is realized within the Work Package 1 (WP1).

The XS-ABILITY project DMP follows the principle of Open Access according to the Horizon Europe guideline summarized in the diagram here below.

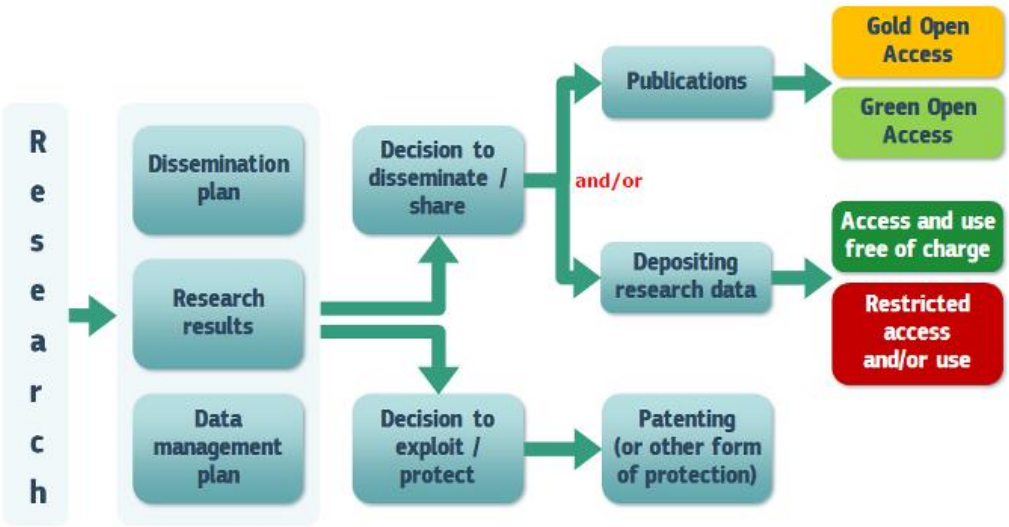

Figure 1: open access to reach data and publication diagram

The other main principles for the XS-ABILITY project DMP are the following:

i. This DMP has been prepared by taking into account the template of the "Guidelines on Data Management in Horizon Europe".

ii. The DMP is an official project Deliverable (D1.2) due in Month 3 (M3 – December 2024), but it will be updated throughout the project. The first initial version will evolve depending on significant changes arising and periodic reviews at relevant reporting stages of the project.

iii. The consortium complies with the requirements of Regulation (EU) 2016/679 and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation - GDPR). Guidance on how these regulations interact with open-access data policy can be found here: https://www.openaire.eu/

iv. Type of data, storage, confidentiality, ownership, management of intellectual property and access: procedures that will be implemented for data collection, storage, access, sharing policies, protection, retention and destruction will be in line with EU standards as described in the Grant Agreement (GA) and the Consortium Agreement (CA).

2.2. XS-ABILITY Data Management strategy

The DMP's Data Management strategy of XS-ABILITY project is focused on the observation of FAIR (Findable, Accessible, Interoperable and Reusable) Data Management Protocols.

This document addresses for each dataset collected, processed and/or generated in the project the following elements:

Dataset reference and name: Internal project Identifier for the dataset to be produced. This will follow the format:

ProjectName_TaskNumber_PartnerName_DataSubset_DatasetName_Version_DateOfStorage, where the project name is XS-ABILITY, the Partner Name represents the name of the data custodian (WP Leader / Task Leader). (ex. *XS-ABILITY_T1.2_CEA_ProgressReport_v2.1_170424*)

Dataset description: description of the data generated or collected, including its origin (in cases where data is collected), nature and scale and to whom it could be useful, and whether it underpins a scientific publication. Information on the existence (or not) of similar data and the potential for integration and reuse.

Standards and metadata: reference to existing suitable standards. If these do not exist, an outline on how and what metadata will be created.

Data sharing: description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling reuse, and definition of whether access will be open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating the type of repository (institutional, standard repository for the discipline, etc.). In cases where the dataset cannot be shared, the reasons for this will be stated (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).

Archiving and preservation (including storage and backup): description of the procedures to be put in place for long-term preservation of the data, including an indication of how long the data should be preserved, the approximate end volume, associated costs, and how these are planned to be covered.

2.3. XS-ABILITY type of data

Annex I provides a list of all datasets currently expected to be generated in the XS-ABILITY project and their planned accessibility. We recognise that this list will develop and grow as the project evolves.

Once generated (or collected), these data will be stored in several formats, which mostly are: Documents, Images, Data.

In particular the following project deliverables are relevant:

D7.1: Dissemination and Communication Plans

This deliverable is part of WP7 and aims at creating and publishing the public communication and dissemination material (website, newsletters, etc.), adapting the support to the target, keeping track of publications and public disclosures.

D7.2: Exploitation plan and IPR management updates for the EC

D7.2 aims at collecting key exploitable results of the project, identifying market opportunities through potential markets or potential customers, and building potential business models using XS-ABILITY solution. Management of Intellectual Property (IP) is also part of this deliverable.

D7.3: Clustering with EU projects, end-users groups and other initiatives

This deliverable will report on activities implemented to foster cross-fertilisation among external projects, standard bodies, regulators, NGOs, networks, communities and other initiatives.

Summarizing, XS-ABILITY generates and collects the following research data relevant for the DMP:

Title	WP No.	Lead Beneficiary	Nature
D7.1: Dissemination and Communication Plans	7	Developia	Report
D7.2: Exploitation plan and IPR management updates for the EC	7	Developia	Report
D7.3: Clustering with EU projects, end-users groups and other initiatives	7	Developia	Report

Table 1: XS-ABILITY research data

Upload instructions - Zenodo
Scientific publications, public deliverables and public datasets must also be uploaded to the *European Commission Funded Research (OpenAIRE) Community* in Zenodo. To do this one must complete the following steps: 1. Create a profile in Zenodo to be able to upload files 2. On the Community site, click the green "New upload" button in the top right corner 3. Enter requested data and confirm the upload. 4. Remember to add the European Commission community in the box labelled "communities". One can use the search function to locate the community and add it. The data will then automatically be uploaded to both communities, so one doesn't have to do it twice. Uploading should be done as soon as possible and at the latest on article publication. Data Controllers are responsible for uploading datasets generated by them. A Data Controller is an individual or organization responsible for determining the purposes and means of processing personal data. Each organization must designate its own Data Controller.

3. FAIR DATA

3.1. Making data findable, including provisions for metadata

Metadata is data on the research data themselves. It enables other researchers to find data in an online repository and is, as such, essential for the reusability of the dataset. By adding rich and detailed metadata, other researchers, can better determine whether the dataset is relevant and useful for their own research. Metadata (type of data, location, etc.) will be uploaded in a standardized form. This metadata will be kept separate from the original raw research data.

The bibliographic metadata include all of the following:

- the terms “European Union (EU)”, “Horizon Europe” and “Euratom”;
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable;
- a persistent identifier.

XS-ABILITY open data will be collected in an open online research data repository: ZENODO. Its repository structure, facilities and management are in compliance with FAIR data principles. Zenodo is an OpenAIRE that allows researchers to deposit both publications and data, providing tools to linking them to these through persistent identifiers and data citations. Zenodo is set up to facilitate the finding, accessing, re-using and interoperating of datasets, which are the basic principles that ORD projects must comply with. Zenodo repository is provided by OpenAIRE and hosted by CERN.

Zenodo is a catch-all repository that enables researchers, scientists, EU projects and institutions to:

- Share research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science;
- Display their research results and get credited by making the research results citable and integrating them into existing reporting lines to funding agencies like the European Commission;

- Easily access and reuse shared research results;
- Integrate their research outputs with the OpenAIRE portal.

Search keywords

Zenodo allows to perform simple and advanced search queries using keywords. Zenodo also provides a user guide with easy-to-understand examples. The Data Controllers at each pilot site will be responsible for uploading public datasets that they have generated and to assign specific keywords relevant to these datasets. Dataset specific keywords must be descriptive to the content of the dataset. In addition, the project has defined a set of general keywords that should apply to all public datasets, scientific publications and public deliverables. These are as follow:

- Embedded nuclear instrumentation
- In-situ characterization
- Robotics
- Artificial intelligence
- Hard-to-access areas
- Nuclear facility
- Autonomous navigation

Naming conventions

Files and folders at data repositories will be versioned and structured by using a name convention consisting as follow:

XS-ABILITY_Origin/PartnerNumber_FileType/Title_YYMMDD.FileExtension (ex. XS-ABILITY_P1_D1.2_241126.doc)

Origin can be:

- WP = WP1, WP2 etc
- EC = European Commission, GA = General Assembly, TC = teleconference
- P1, P2, P3 etc = partner number

FileType are:

- D stands for Deliverable

- DS stands for DataSet
- F stands for File (generic, ex. images, table, document)

Title: agenda, minutes, slides, etc

Version numbers

Individual file names will contain version numbers that will be incremented at each revision (V*zz*). Zenodo provides DOI versioning of all datasets uploaded to their communities, which allows us to edit and update the uploaded datasets after they have been published. This also allows us to cite specific versions of an upload and cite all versions of an upload.

3.2. Making data openly accessible

In order to maximise the impact of XS-ABILITY research data, the results are shared within and beyond the consortium. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through open access data repositories.

The XS-ABILITY project datasets are first stored and organized in a database by the data owners (personal computer, or on the institutional secure server) and on the project database (project website). All data are made available for verification and re-use, unless the task leader can justify why data cannot be made openly accessible. To protect the copyright of the project knowledge, Creative Commons license will be used in some cases. To use the Creative Commons license, include a clear attribution to the original source, specify the type of license (e.g., CC BY-NC-SA 4.0), and follow the conditions outlined, such as non-commercial use or sharing alike. Ensure this information is recorded in the metadata of digital datasets or visibly on physical copies. It is also recommended to register the use of the license with the Creative Commons organization through their website to ensure proper documentation and visibility.

The XS-ABILITY dataset deliverables are both public (data access policy unrestricted) and they will be accessible by:

- XS-ABILITY project web site

- Partners database
- OpenAIRE
- ZENODO (https://zenodo.org) for ORDP data and datasets
- Open access journals

All data deposited on Zenodo are accessible without restriction for public. For other data, potential users must contact the IPR (Intellectual Property Rights) team or the data owner in order to gain access. If necessary, appropriate IPR procedure (such as non- disclosure agreement - NDA) will be used

3.3. Making data interoperable

Partners will observe OpenAIRE guidelines for online interoperability, including OpenAIRE Guidelines for Literature Repositories, OpenAIRE Guidelines for Data Archives, OpenAIRE Guidelines for CRIS Managers based on CERIF-XML. These guidelines can be found at: https://guidelines.openaire.eu/en/latest/.

Partners will also ensure that XS-ABILITY data observes FAIR data principles under Horizon Europe open-access policy.

In order to ensure the interoperability, all datasets will use the same standards for data and metadata capture/creation.

As the project progresses and data is identified and collected, further information on making data interoperable will be outlined in subsequent versions of the DMP. Specifically, further details will be given on the vocabularies, standards, or methodologies for data and metadata to ensure interoperability. This includes whether a standard vocabulary is used for all data types to support interdisciplinary interoperability.

3.4. Increase data re-use

Creative Common (CC) Licensing will be used to protect the ownership of the datasets. Different types of licenses will be considered.

However, an embargo period (usually 12 months for targeted journals) may be applied if the data (or parts of data) are used in published articles in “Green” open access journals.

For datasets deposited on a public data repository (i.e. Zenodo) the access is unlimited.

Restrictions on re-use policy are applied for all protected data (see Figure 1: Open access to research data and publication decision diagram), which’s re-use will be limited within the project partners.

Other restrictions could include:

- the “embargo” period imposed by journals publication policy (Green Open access);
- some or all of the following restrictions may be applied with Creative Commons licensing of the dataset (attribution, non-commercial, share-alike, etc.).

Internal process of Quality evaluation is activated throughout the entire project duration to assess both project data /products and project process (See the D1.1 - Project Handbook). An internal peer review is performed for the main project deliverables to guarantee the deliverable is developed with a high level of quality. It is recommended that WP leaders make sure that an internal reviewer (in their team or in the consortium) checks the quality of the documents produced.

Longevity

For data published in scientific journals, the underlying data will be made available no later than by journal publication. The data will be linked to the publication. Data associated with public deliverables will be shared once the deliverable has been approved and accepted by the European Commission. For other public datasets not directly linked to a scientific publication or deliverable, such datasets will be made available upon assessment by the Data Controllers that it is ready for publishing, and in the final month of the project at the latest.

Open data can be reused in accordance with the Creative Commons licences. Data classified as confidential will as default not be reusable due to privacy concerns.

The public data will remain reusable via Zenodo for at least 20 years. This is currently the lifetime stated by the host laboratory CERN. In the event that Zenodo has to close their operations, they

have provided a guarantee that they will migrate all content (including metadata) to other suitable repositories.

3.5. DMP review process & data inventory

Internal process of quality evaluation and reporting is activated throughout the entire project duration to assess both project data /products and project process (See the D1.1 - Project Handbook).

Results data will be also analysed and collected throughout the project entire duration. To this purpose follow-up of communication and dissemination actions will be collected by the leader of WP7 and presented at EB meetings. It includes the description of articles, papers and scientific publications too. Thus, all research data generated and publications will be analysed and described by using the Data Inventory Table (Annex I), WP leaders and partners authors of publications are required to fill in periodically.

Further updating of the DMP will include the eventual updating of online research data repository where data are collected and shared; and the data, the description of dataset, and research data, gradually generated and collected.

4. ALLOCATION OF RESOURCES

Costs related to open-access to research data in Horizon Europe are eligible for reimbursement under the conditions defined in the Horizon Europe GA, but also other articles relevant for the cost category chosen. Project beneficiaries will be responsible for applying for reimbursement for costs related to making data accessible to others beyond the consortium.

The costs for making data FAIR includes:

- Fees associated with the publication of scientific articles containing project's research data in "Gold" Open access journals. The cost sharing, in case of multiple authors, shall be decided among the authors on a case-by-case basis;
- Project Website operation: to be determined;
- Data archiving at Zenodo and on other online data base: free of charge;
- Copyright licensing with Creative Commons: free of charge.

The project members of the General Assembly are also responsible of the Data Management of XS-ABILITY dataset and research data in accordance with each organization internal Data Protection Officer (DPO).

Each partner is responsible for the data they produce. Any fee incurred for Open Access through scientific publication of the data will be the responsibility of the data owner (authors) partner(s).

The overall responsibility for data management processes within the project lies with the project coordinator, Maugan Michel from CEA. In Horizon Europe projects, the coordinator ensures that data management procedures are properly implemented and followed. However, individual partners retain responsibility for the data they produce, and any fees incurred for Open Access publication of their data.

5. DATA SECURITY

The following guidelines, which are aligned with common best practices, will be followed in order to ensure the security of the data:

- Store data in at least two separate locations to avoid loss of data;
- Encrypt data if it is deemed necessary by the participating researchers;
- Limit the use of USB flash drives;
- Label files in a systematically structured way in order to ensure the coherence of the final dataset.

All project deliverables and data will be stored and shared in the Project SharePoint restricted to the project consortium. As an initial step, only the Consortium Partners will have access to the repository where dataset and metadata are filed.

Following, scientific publications and articles, the dataset deliverables and the final demonstrator research results will be shared through Zenodo and other database to promote the “FAIR” data making

6. ETHICAL ASPECTS

The WP1 aims at ensuring that ethical requirements are met for all research undertaken in the project, including data management aspects, in compliance with Horizon Europe ethical standards.

All partners will assure that the EU standards regarding ethics and data management are fulfilled in compliance with the ethical principles and confidentiality (as set out in the GA).

In addition:

1. In accordance with the GDPR 2016/679, the data controllers and processors are fully accountable for the data processing operations.
2. Templates for informed consent forms and information sheet are also available. More details in relation to Ethics (and Security) in relation to Data Management can be found in Section 5 of the GA.

The project data will be exchanged between the partners of the XS-ABILITY consortium at all times during the project.

7. CONCLUSION

This document describes the main principles and guidelines for the Data Management for the XS-ABILITY project. As a live document, it will be updated throughout the project lifetime. Further updating of the DMP will include the eventual updating of online research data repository where data are collected and shared and the data the description of dataset and research data gradually generated and collected.

8. REFERENCES

European Commission. *Horizon Europe Data Management Plan Template.*

European Commission. Horizon Europe Data Management Plan Template

European Commission. Horizon Europe. Guidance for the classification of information in research project[1]

European Commission. Horizon Europe. Guidelines on FAIR Data Management in Horizon Europe[2]

[1] classification-of-information-in-he-projects_he_en.pdf (europa.eu) ; how-to-handle-security-sensitive-projects_en.pdf (europa.eu)

[2] programme-guide_horizon_en.pdf (europa.eu) ; programme-guide_horizon_en.pdf (europa.eu) ; h2020-hi-oa-data-mgt_en.pdf (europa.eu) ; Open Data, Software and Code Guidelines | Open Research Europe (europa.eu)

9. ANNEX I – DATASETS COLLECTED IN THE PROJECT

Dataset reference and name: Project management documentation Person in charge: Maugan Michel		
Data Type	**Data Standards - Formats**	**Data Generation Software**
Reports, documentation and presentations	.docx, .xlsx, .pdf, .pptx	Microsoft word, excel, powerpoint, adobe acrobat reader
Estimated Data Size	**Data Sharing**	**Data Storage and Preservation**
5 Go	With European Commission, with the consortium, specific deliverables with large public	Microsoft Teams

Data validation? ☑ Yes ☐No

All data linked to a lab book reference? ☐ Yes ☑ No

Dataset reference and name: xxxx Person in charge: xxx		
Data Type	**Data Standards - Formats**	**Data Generation Software**
Estimated Data Size	**Data Sharing**	**Data Storage and Preservation**

Data validation? ☑ Yes ☐No

All data linked to a lab book reference? ¨ Yes þ No

Dataset reference and name: Prescription data Person in charge: xxx (IBV)		
Data Type	**Data Standards - Formats**	**Data Generation Software**
Estimated Data Size	**Data Sharing**	**Data Storage and Preservation**

Data validation? ☑ Yes ☐No

All data linked to a lab book reference? ¨ Yes þ No

=== End of the document ===