DELIVERABLE D4.2 –
Data Management Plan – V2

COLOSSE
Central European Platform for Plasma-Enabled Surface Engineering

Project Acronym	COLOSSE
Project Full Name	Central European Platform for Plasma-Enabled Surface Engineering
Grant Agreement No.	101158464
Call	HORIZON-WIDERA-2023-ACCESS-04
Type of action	HORIZON Coordination and Support Actions
Project starting date	1 April 2024
Project duration	36 months
Deliverable number	D4.2
Deliverable name	Data Management Plan V2
Work Package	WP4
Type	R — Document, report
Main author	E. Skalická
Lead beneficiary	Masaryk University
Dissemination Level	PU - Public
Due Date	June 2025
Date	28. 06. 2025
Version	2.0

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101158464. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

COLOSSE
Central European Platform for
Plasma-Enabled Surface Engineering

Table of contents

Abbreviations

CEPLANT - R&D centre for plasma and nanotechnology surface modifications

DECP - Dissemination, Exploitation, Communication Plan

DMP – Data Management Plan

EOSC – European Open Science Cloud

ERDF - European Regional Development Fund

EU – European Union

FP – Framework programme

HE – Horizon Europe

HRS4R - Human Resources Strategy for Researchers

MU – Masaryk University

NDI - National Data Infrastructure

NRP - National Repository Platform

R&I – Research and innovation

WP – Work package

Executive summary

The project is designed to achieve its objectives through four key strategies: enhancing connections with world-leading research and innovation (R&I) centers, fostering internationalization of human resources, developing interdisciplinary skills, and ensuring the sustainability of the partnership. These efforts will involve creating collaboration strategies, implementing international recruitment schemes, providing targeted training for researchers, and developing concepts for future funding opportunities. The Data Management Plan (DMP) will ensure that all data generated throughout the project is properly managed, stored, shared, and preserved, adhering to best practices and facilitating long-term access for relevant stakeholders.

Key words

Data, management, FAIR principles, storage, collection, formats

Summary of updates

This document is the second version of the Data management plan. It includes minor updates such as adding of a link to Zenodo and the project website, Furthermore, the information regarding the publication date of deliverables has been added for clarification.

1 Introduction

1.1. Project in brief

COLOSSE project aims to achieve its goals through four specific objectives:

1. Strengthen the connections of COLOSSE centres to world-leading R&I centres. (implemented through WP1)
   a) Develop joint Internationalization Strategy of COLOSSE.
   b) Define collaboration areas with strategic partners.
   c) Exchange staff with world-leading R&I counterparts.
   d) Promote COLOSSE facilities as excellent and reliable partners for R&I.

2. Build conditions that will enable internationalization of human resources. (implemented through WP2)
   a) Develop joint HR Strategy that will build upon HRS4R of COLOSSE centres, with special focus on international recruitment, on-boarding, and continuous support for incoming staff.
   b) Implement pilot international recruitment scheme.

3. Develop skill-set that enables interdisciplinary and intersectoral collaboration and facilitates creativity. (implemented through WP2)
   a) Train soft and transferable skills of researchers based on needs identified in the joint HR Strategy.
   b) Complement the set of researchers' hard skills through participation in training workshops and schools internationally, based on their individual career development plans.
   c) Facilitate development of young researchers' network connecting the COLOSSE centres.

4. Enable sustainability of the COLOSSE partnership through synergic use ERDF and HE/FP resources. (implemented through WP3)
   a) Train research support personnel on the pre- and post-award aspects of HE/FP projects.
   b) Develop project concepts and consortium cores to prepare for upcoming HE/FP funding opportunities.

These objectives will generate various types of data, including collaboration strategies, human resource strategies, training materials, and project concepts. The Data Management Plan (DMP) will outline how this data will be collected, stored, shared, and preserved, ensuring compliance with relevant standards and enabling long-term accessibility for all stakeholders.

The project will have a major impact on academic environment and culture of research in the COLOSSE centres, and thus we intend to shape some of our actions so that they support Open Science and FAIR Data Management:

- All participants of COLOSSE R&I mobility scheme (T1.2) will be requested to publish their results in compliance with open access rules of Horizon Europe rules (CC-BY).
- The postdocs recruited under T2.2 will be requested to elaborate a data management plan concerning their research in line with the FAIR principles; their research results will also be published in open access according to Horizon Europe rules.
- The career development framework prepared in Human resources development strategy (T2.1) will specifically focus on recognition of the diversity of research outputs including publication of data and open science practices.
- Our primary aim of opening up the ERDF-funded facilities to collaboration will be guided by principles of open access to research infrastructures (in CEPLANT, this is further facilitated by national funding).
- All project deliverables will be public.

1.2. Organisation of the COLOSSE project

The project is organized in five Work packages:

WP1	Objectives: Reinforce links with strategic partners. Develop COLOSSE Internationalization Strategy. Establish close R&I interactions with strategic partners through mobility of researchers in all career stages
WP2	Objectives: Implement HRS4R strategies at COLOSSE centres through adoption of joint HR strategy

WP3	Objectives: Develop competences needed to prepare winning proposals among researchers and research managers and administrators. Identify topics and ideas with potential for development of projects suitable for HE/FP applications. Develop ideas into full-scale project concepts and project proposals.
WP4	Objectives: Ensure efficient project management, monitoring, reporting and data management.
WP5	Objectives: Design and implement the Dissemination, Exploitation, Communication Plan (DECP).

1.3. Management

The coordinator will collaborate closely with the relevant professional services at Masaryk University, including the MUNI Open Science team and the MUNI Personal Data Protection Office.
For processing personal data (surveys and feedbacks from trainings), we will apply standard procedures at Masaryk University that are accessible here: https://www.muni.cz/en/about-us/official-notice-board/personal-data-protection.

The Data Management Plan (DMP) is a living document. This is the second version, with the next update planned for the end of the project (M36) or whenever a significant change occurs.

From 2024, there will be a gradual introduction of services and components of the so-called National Data Infrastructure (NDI), as envisaged by the Architecture for the implementation of EOSC in the Czech Republic, to manage research data in the Czech Republic. In particular, the use of the so-called National Repository Platform (NRP) component is then discussed. Due to the fact that this infrastructure is just being created, it is currently not possible to provide more detailed information than what is provided at www.eosc.cz. The project consortium is aware of this initiative and will continuously monitor it in order to evaluate the possibility of using the upcoming services and storage capacities of the national solution.

2 Data Summary

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

The project is a coordination and support action, and thus will not directly generate any research results or data. The DMP will focus on coordination and support measures – data of participants to training and mobility, training materials, R&I priorities, strategies, guideline for selection procedure of postdoctoral researchers etc.

What types and formats of data will the project generate or re-use?

In order to enable data sharing, we will be using well-known standard formats and types as follows: tables (XLSX), documents (DOCX, PDF), presentations (PPTX, PDF), images (JPG, PNG, etc.), video file formats (MP4, WMV, etc.). These formats are suitable for long-term archiving purposes.

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

Based on the data generated during the project, we will achieve the goal of the project – to increase the participation of Czech and Slovak plasma-enabled surface engineering R&I centres in Horizon Europe and future EU Framework Programmes for R&.

What is the expected size of the data that you intend to generate or re-use?

We will have only a small amount of data stored (GiBs).

What is the origin/provenance of the data, either generated or re-used?

Data generated from trainings, feedbacks, strategies, proposal drafts

To whom might your data be useful ('data utility'), outside your project?

The data generated within the project might be useful to the scientific community, companies, the public, and definitely project partners (researchers, students and administrative staff). F.e. the HR strategy can serve as a good practice for effective onboarding and recruitment of international staff. Also the training strategy can be easily transfer to other institutions to enhance soft and transferable skills.

3 FAIR data

3.1. Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

A persistent identifier will be added to the data (documents) when placed into the long-term repository (e.g., Zenodo). However, the project consortium will discuss the option of assigning the identifier to all deliverables directly by MU institution.

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Most metadata will be made available through the project deliverables as well as through the project dissemination channels (e.g., project website and social media accounts). For the results deposited in Zenodo, it will be possible to export their metadata in standard formats such as DataCite XML.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Keywords will be provided within the description of each deliverable as well as executive summary.

The standard format of file naming will be:

Number of deliverable_name_version (f.e D4.1_Data Management Plan_V1)

Will metadata be offered in such a way that it can be harvested and indexed?

Metadata will be available in a form that can be harvested and indexed (managed by the used repository/repositories). Zenodo issues data in a standard way via OAI-PHM.

3.2. Making data accessible

We will use institutional data storage (institutional Microsoft M365 cloud space) at Masaryk University for storing data and outcomes. It is a secured institutional cloud service with adequate legal agreements, professional technical maintenance and operations, and sufficient access control. The access for project partners is possible.

This storage type is higly suitable for storage and collaborative processing of the data types produced in our project.

Repository

Will the data be deposited in a trusted repository?

Zenodo is planned to be utilized. The project website will also serve as a repository for project outputs.

Link to COLOSSE Zenodo: https://zenodo.org/communities/colosse/

Link to COLOSSE website: https://colosse.eu/

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Yes, Zenodo seems to be a suitable solution.

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

Yes, the Zenodo repository allows assignment of DOIs to the records.

Data

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

All deliverables produced will be made public as soon as possible (after approval by the Project Officer). They will be accessible on the project website and Zenodo.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

N/A

Will the data be accessible through a free and standardized access protocol?

Data will be openly available on the project website or Zenodo repository via standard protocols like HTTPS.

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

N/A

How will the identity of the person accessing the data be ascertained?

N/A

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

N/A

Metadata

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

All our metadata can become completely open over time (website, Zenodo). CC-BY licensing will be applied.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

All repositories provide long-term storage of data. Data will be available for at least 5 years after the end of the project.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

No specific software is needed. The documents will be exported in pdf file to ensure the accessibility.

3.3. Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

We will use standard open data formats that are common and easily usable to allow inter-disciplinary interoperability.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

N/A

Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

We will include references to other data using PIDs like DOI where appropriate.

3.4. Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

All the used sources will be properly cited and used methodology will be part of the documents.

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

All data will be available on the project website and in trusted repositories (see above). Our data will be licensed using free licenses.

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Data produced and/or used in the project will be useable by third parties.

Will the provenance of the data be thoroughly documented using the appropriate standards?

The versioning of the document and the post-implementation of the work will be recorded directly in the documents.

Describe all relevant data quality assurance processes.

Before the document is published or submitted as an official project output, it will be reviewed and approved by all partners. Any feedback or proposed changes should be considered by the original authors, who will decide whether to implement or reject them. The final review will be conducted by the Project Manager of the coordinator.

4 Other research outputs

There will be no other project outputs.

5 Allocation of resources

FAIR is a central part of our data management; it is considered at every decision in our data management plan. We use the FAIR data process ourselves to make our use of the data as efficient as possible.

We do not require any hardware or software in addition to what is usually available in the institute.

The Project Manager of the coordinator, in collaboration with the work package leaders, is responsible for data management. To implement the DMP, additional specialist expertise from MU was utilized, as we have trained support staff available for this purpose.

The postdocs recruited under WP2 will be requested to elaborate a data management plan concerning their research in line with the FAIR principles; their research results will also be published in open access according to Horizon Europe rules.

6 Data security

We are running the project in a collaboration between different groups and institutes. A collaboration agreement that describes who can have access to what data in the project is set.

All data will be stored on the institutional storage infrastructures and automatically backed up.

7 Ethics

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

According to the Grant Agreement, any research connected to the project will be carried out in compliance with fundamental ethical principles. If we identify an ethical issue, we will approach an ethical committee at a relevant institution to get approval and proceed according to standard procedures and regulations.

Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

We will follow General Data Protection Regulation standard procedures as mentioned in section 1.3.

8 Other issues

We will be using the following policies and procedures for data management:

MU Open Science Strategy 2022–2028 (https://openscience.muni.cz/media/3477939/en-a-strategicka-cast-strategie-open-science-mu-2022-2028-uprprebal.pdf). Strategic framework for effective implementation of Open Science at Masaryk University (Coordinator).