



Funded by
the European Union



InPlasTwin

Increasing expertise in micro- and nanoplastics analysis through twinning action

Horizon Europe Grant agreement ID: 101160289

D 11.1 Data management plan

inplastwin.eu



Author(s) / Organisation(s)	Jožef Stefan Institute - JSI
Contributor(s)	Janja Vidmar
Work Package Title	Project Management and Coordination (Phase 1)
Linked task title	Task 11.3 Data management
Partner responsible	Jožef Stefan Institute
Actual Delivery Date	31. 3. 2025
Dissemination Level	PU - Public
Abstract:	The present report entitled "Data management plan" (DMP) summarizes the InPlasTwin activities and actions towards comprehensive data management strategy that is aligned with EC guidelines and FAIR Data principles. DMP defines procedures for the acquisition, creation, storage, use, archiving, and deletion of all project-related data, metadata, and information. This document represents the initial iteration of the DMP (v1.0) and will be updated regularly to track its implementation.

Document Revision History			
Date	Version	Author/Contributor/ Reviewer	Summary of main changes
17. 2. 2025	V1	Janja Vidmar	1 ST Draft
31. 3. 2025	V2	Janja Vidmar	Final draft – after internal revision

Participant Number	Participant organisation name	Short name	Country
1.	INSTITUT JOZEF STEFAN	JSI	Slovenia
2.	GEOPONIKO PANEPISTIMION ATHINON	AUA	Greece
3.	UDRUZENJE ZA PREDUZETNISTVO I INOVACIJE FOODSCALE HUB	FSH	Serbia
4.	VLAAMSE INSTELLING VOOR TECHNOLOGISCH ONDERZOEK N.V.	VITO	Belgium
5.	HAVFORSKNINGSINSTITUTTET	IMR	Norway
6.	DANMARKS TEKNISKE UNIVERSITET	DTU	Denmark

Disclaimer



Funded by the European Union. 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 (REA). Neither the European Union nor the granting authority can be held responsible for them.

Title: Increasing expertise in micro- and nanoplastics analysis through twinning action

Type of action: HORIZON Coordination and Supporting Action

Topic: HORIZON-WIDERA-2023-ACCESS-02-01

Start date of project: 1 October 2024 • Duration: 36 months

© InPlasTwin Consortium, 2024

Reproduction is authorised provided the source is acknowledged.

Executive summary

The present document is the deliverable D11.1 “Data management plan” of the InPlasTwin project, funded by the European Commission’s under its Horizon EUROPE Framework Programme for Widening participation and spreading excellence under the grant agreement no. 101160289.

The main objective of this deliverable is to address all the requirements and practices for managing the project’s data and documentation throughout the data life cycle.

This document describes the data management of the InPlasTwin project and outlines: what data will be collected, processed and generated throughout the project; how the research data will be handled during and after the end of the project; which methodologies and standards will be applied; whether and how data will be shared and made accessible in open access; how the data will be maintained, preserved and curated, both during and after the end of the project.

This deliverable is considered a dynamic living document that will be regularly updated over the course of the project as the new types of data emerge as a result of the project activities.

Table of Contents

1.	Introduction	7
1.1.	Project summary	7
1.2.	Document scope	7
2.	Data summary	8
2.1.	Purpose of data generation and its link to project objectives	8
2.2.	Data types and data description	9
2.3.	Re-use of existing data	10
2.4.	Expected data size and origin	10
2.5.	Data utility	10
3.	Fair data principles	10
3.1.	Making data findable, including provisions for metadata	10
3.2.	Making data accessible	11
3.2.1.	Repository	11
3.2.2.	Data	12
3.2.3.	Metadata	12
3.3.	Making data interoperable	13
3.4.	Increasing data re-use	13
4.	Other research outputs	14
5.	Allocation of resources	14
6.	Data security	15
7.	Ethics	15

List of Abbreviations and Acronyms

ASCII	American Standard Code for Information Interchange
AUA	Agricultural University of Athens
CC licences	Creative Commons licences
DMP	Data management plan
DOI	Digital Object Identifier
DTU	Technical University of Denmark
EC	European Commission
EOSC	European Open Science Cloud
EU	European Union
FAIR	Findable, accessible, interoperable, and reusable
FSH	Foodscale Hub
GDPR	General Data Protection Regulation
HE	Horizon Europe
IMR	Institute of Marine Research
IP	Intellectual Property
JSI	Jožef Stefan Institute
MNPs	Micro- and nanoplastics
PCCS	Provenance and Context Content Standard
SO	Specific Objective
VITO	Flemish Institute for Technological Research
WP	Work Package

1. Introduction

1.1. Project summary

The InPlasTwin project aims to enhance the research and innovation capabilities of the Jožef Stefan Institute (JSI) and the Agricultural University of Athens (AUA) in micro- and nanoplastics (MNPs) analysis by collaborating with leading European institutions like Flemish Institute for Technological Research (VITO), Institute of Marine Research (IMR), and Technical University of Denmark (DTU). InPlasTwin will provide access to advanced equipment, training, and expertise in MNP detection, extraction, quantification, identification, characterization, and plastic additive analysis to ensure the achievement of its objectives. The partnership with Foodscale Hub (FSH) will expand the reach of JSI and AUA's research, raise awareness about plastic pollution in the agri-food sector, and promote diversity and ethical practices in research. Moreover, InPlasTwin activities will focus on improving management and administrative skills at JSI, the project coordinator institution.

Through twinning activities such as researcher exchanges, training workshops, and scientific meetings, InPlasTwin aims to transfer knowledge and skills, foster collaboration, and disseminate research findings. By advancing MNP research and promoting responsible research practices, InPlasTwin will contribute to improved environmental sustainability and food safety.

1.2. Document scope

This document, the Data Management Plan (DMP), provides all relevant details about specific characteristics of the InPlasTwin project in terms of data management. Within this document, the preliminary information on data that will be gathered, generated, or processed is described, the standards that will be used for its management, the ways the data will be exploited and shared, and how the data will be preserved during the InPlasTwin project lifetime and upon its completion with details such as ethical, privacy, and security issues. This deliverable shows how the partners involved in this project are going to ensure that research data are findable, accessible, interoperable, and reusable – FAIR.

The document structure is aligned with the template recommended for Horizon Europe beneficiaries (EU Grants: Data management plan [HE]: V1.1 – 01.04.2022), which includes the following key elements: Data summary, FAIR data principles that aim to make data easily findable, accessible, interoperable, and reusable, Other research outputs, Allocation of resources, Data security, and Ethics.

The DMP deals with the following questions:

- ✓ What is the purpose of the data collection/generation and its relation to the objectives of the project?
- ✓ What types and formats of data will the project generate/collect?
- ✓ Will the project reuse any existing data, and how?
- ✓ What is the origin of the data in this project?
- ✓ What is the expected size of the data?
- ✓ To whom might the data be useful ('data utility')?
- ✓ How will the FAIR principles be implemented within the project?
- ✓ Which resources are dedicated to the FAIR principles' implementation?
- ✓ Which data security measures are set?
- ✓ What are the ethical concerns regarding data management?

It is important to note that DMP in its present form is a guiding document established early in the project. As such, it cannot cover all possible types of data and outcomes of the research activities. The document will be regularly updated as the project activities progress and will always rely on the provisions set in the Grant agreement and Consortium Agreement. We will also utilize ARGOS-OpenAire for the management, validation, monitoring, and maintenance of the DMP. This tool will enable us to evaluate the implementation of Open Science practices throughout the entire life cycle of the project.

2. Data summary

2.1. Purpose of data generation and its link to project objectives

The project's main objective is to strengthen the research and innovation capacity of the JSI and AUA in the analysis of micro- and nanoplastics in the environment and food through networking, knowledge sharing, and research activities with leading research institutions in Europe.

The DMP for the InPlasTwin project is created in line with its specific objectives, which are summarised below:

SO1: Enhancing long-term and sustainable networking activities of the JSI and AUA with project partner institutions (FSH, VITO, DTU, IMR) through knowledge transfer and exchange of best practices in the investigation of MNPs in the environment and food;

SO2: Strengthening the scientific and innovation capacity of JSI and AUA in investigating MNPs in the environment and food;

SO3: Strengthening the research profile of the JSI and AUA and their staff, including increasing the management and administrative skills of the JSI and diversity inclusion and ethic practices at JSI, AUA, and FSH;

SO4: Enlarging the social impact of research and raising awareness of different target groups on the environmental and food impacts of MNPs.

To support the realisation of all these objectives, the following data will be generated throughout the project implementation:

- **Training tools and educational materials** will be developed to support knowledge transfer and the exchange of best practices from DTU, VITO, and IMR to JSI and AUA. These materials will focus on the knowledge exchange on the generation of MNPs in agricultural contexts, analytical techniques for the detection of MNPs, and the analysis of plastic additives (SO1), as well as on improving management and administrative skills, promoting diversity inclusion and ethical practices at partner institutions from widening countries (SO3).
- **Research data and research results** will be generated through scientific and innovation capacity-building activities (SO2) planned under WP7 – Boosting research on the plant uptake of MNPs, and WP8 – Validating scientific and innovation capacity, both of which include research component. In WP7, the DMP covers the research data generated from a lab-scale experiment designed to deepen the understanding of MNP uptake and its effects on plants. In WP8, the DMP addresses research data from a small-scale field experiment designed to assess the food safety and quality of strawberries grown using mulching films.
- **Project dissemination materials** will promote the project and share its results with the widest possible audience at national, regional, and international levels, ensuring the project's visibility extends beyond its duration (SO4).

Considering that this project will generate and collect large amount of (research) data coming from different training materials and analytical techniques, the data collection, document sharing, and collaboration among partners is of outmost importance to ensure adherence to best practices in the acquisition, creation, storage, use, archiving and deletion of any data and information obtained or generated during the project.

InPlasTwin is committed to implementing Open Science practices to enhance transparency, collaboration, and accessibility, thereby maximizing the societal impact of our research and contributing to the advancement of scientific knowledge. Emphasis will also be placed on FAIR data management.

2.2. Data types and data description

The InPlasTwin project plans to carry out training, education, research, dissemination, and communication actions across various disciplines involving multiple layers of data collection and data types.

To support the realisation of all these activities, the data management plan should consider (at least) three types of data:

- **Training tools and educational materials** will consist of documents, presentations, webinars, hands-on and online training, etc., produced within the knowledge-exchange activities in WP3, WP4, WP5, and WP6.
- **Research data and research results** generated or collected throughout the project will include sample information, raw laboratory data (laboratory experiments, protocols, results), quantitative tabular data with minimal metadata (analytical performance data, laboratory experimental data), digital images related to the project research activities (e.g., experimental data presentation, maps), etc. The workflow from samples to raw data and analysis tables/graphs consists of the following steps:
 - **Sample metadata** – Information about the sample (e.g., collection time and location, sample type, etc.)
 - **Preparation metadata / Experimental metadata** - Information about the standard operating procedures used for sample preparation and conducting experiments (e.g., extraction protocols, chemicals, instrumentation, and instrumental parameters), including where the raw sequence datasets are saved.
 - **Raw metadata** - Primary large-scale experimental data types collected, such as spectral datasets, mass spectra datasets, images & elemental maps obtained using different analytical techniques for MNPs detection.
 - **Data processing metadata** – Information on the processing and analysis of raw experimental data from each sample, utilizing software integrated within the instrumentation or other specialized software available.
 - **Results metadata** – Data on the features detected in the samples, such as particle size distribution, particle mass, particle number concentration, and polymer composition.
- **Project dissemination materials** will consist of web and/or press articles, digital video data (news and media items), presentations, webinars, promotional materials, etc.

Scientific Publications produced by any scientific and/or innovation research sector of the project will be published under Full Open Access principles, adhering to transparent peer review standards. Additionally, we will utilize Open Research Europe, sharing preprints via repositories such as Zenodo

(an OpenAIRE-compliant repository). Moreover, protocols, workflows, and analytic scripts will be shared in repositories such as protocols.io and GitHub in compliance with FAIR principles.

InPlasTwin will also produce Project report deliverables, as described in its Description of Action. Those reports are of several confidentiality levels, yet most fall under the “Public” access class. As such, they shall be published in open access repository and the EC project management platform, as supported by the relevant software. Confidential reports shall be only uploaded in the EC project management platform and protected appropriately by it.

The project will collect raw data in ASCII format, processed data in XLSX or CSV format, text in DOC/DOCX or PDF format, project slideshows in PPT/PPTX format, and images in PNG/JPEG/TIFF formats.

2.3. Re-use of existing data

Most of the research data will be collected and/or generated during the project (new data) and will be based on setups and test conditions specific to the InPlasTwin project.

However, certain project activities, particularly those related to training and education, will build upon internal data from the participants (e.g., knowledge of protocols already established at partner institutions giving training), which has been generated by project partners in the past. This data will be made available for secondary analysis and reuse, subject to intellectual property rights. Additionally, the project may extract and organize publicly available data from research articles published in scientific journals through systematic reviews and meta-analysis. Some software for specific analytical techniques may also include spectra and data from existing literature, and technical reports will be consulted during the method development phase.

2.4. Expected data size and origin

As the project progresses, the expected size of the data is to be evaluated.

The data generated in the InPlasTwin project will come from various sources, including training tools and educational materials developed or reused by project partners during knowledge-exchange activities. The majority of the research data will be derived from various analytical measurements using instruments available in the laboratories of the project partners.

2.5. Data utility

The data generated in the InPlasTwin project will be suitable for use by scientific community, research institutions and academia focused on the project topics related to the implications of MNPs in the environment and food. It will also be useful for farmers, food processors and manufacturers, industry and technology experts, the general public, and policymakers at different levels.

3. Fair data principles

In line with responsible data management and open access principles, the project will adhere to the FAIR principles (Findable, Accessible, Interoperable, and Reusable) and promote open data exchange as the standard for EU-funded research.

3.1. Making data findable, including provisions for metadata

Data to be shared with the research community and/or specific stakeholders for which the data will be useful should be easy to find (by both human and machine) and therefore will have to meet the

following criteria: 1) Each data type will use an internationally accepted file format for its effective sharing, reuse, and preservation, which will improve interoperability of data between different systems. 2) Data will be accompanied with relevant documentation that will describe context, content and structure of data. That will include all the important information on data origin, collection methods, and any manipulation or processing applied to data.

Persistent identifiers will be applied so that datasets can be reliably and efficiently located and referred to, as well as cited and re-used. Data sets will be structured following established data processing principles and annotated using ontologies. The Dublin Core metadata schema will be utilized for data storage, enhancing discoverability and reusability. A Digital Object Identifier (DOI) will be used to assign a unique name (rather than a location) to an entity within digital networks. This ensures a system for persistent, actionable identification and the interoperable exchange of managed information on digital platforms. A DOI, issued by the International DOI Foundation, is a type of Persistent Identifier (PID) that provides a permanent link to a digital object, allowing reliable reference even if its location or metadata changes over time. The format for raw data will follow the structure: [YYMMDD]-[project number]-[Affiliation]-[WP#]-[T###]-[version nr]-[keyword]-[keyword].

Specific types of metadata that will be created for digital objects depend on the needs and priorities of the domain and community in which they are being used. Rich and standardised metadata will be created as it is essential for enabling discovery, understanding, and reuse of digital objects. These include descriptive metadata (title, author, date subject), which provides information about the creator of a resource, and its content; structural metadata (file format, size), which details how data elements are organized, their relationships, and their structure; administrative metadata, which gives information about the resource's origin, type, access rights, preservation strategy, and history; and preservation metadata (migration, backup, recovery), which focuses on maintaining the resource's longevity.

In alignment with Horizon Europe's open science policy outlined in the Programme Regulation (Art. 10 and 35(3)) and the General Model Grant Agreement (Art. 17), we will adopt Creative Commons Attribution International Public License (CC BY) for metadata. This approach ensures the protection of knowledge and legitimate commercial interests while facilitating the free circulation of documents and preserving authors' and the project's intellectual property rights (CC BY-NC, CC BY-ND).

To optimize the possibilities for discovery and potential re-use of data, a table of search keywords will be established, including the name of the project (InPlasTwin) as well as other keywords relatable to the project's topic (e.g., microplastics, agriculture, mulching films, etc.).

3.2. Making data accessible

3.2.1. Repository

When sharing results, we will follow the principle "as open as possible, as closed as necessary", and data that can be shared will be shared under Creative Commons licenses.

As per the project grant agreement, each beneficiary must ensure open access to all peer-reviewed scientific publications relating to its results. Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications, as well as any other data generated during the project. Research data will be deposited to the Zenodo repository unless a decision has been made to protect the results. All the data deposited in Zenodo will be made accessible to any user, free of charge, using the appropriate license. Zenodo.org is open, free, searchable, and structured with flexible licensing, allowing for storing all types of data: datasets, images, presentations, publications and software. In addition, Zenodo allows

researchers to deposit both publications and data while providing tools to link them. A DOI will be assigned to all datasets when it is uploaded to Zenodo. This DOI can be used in all related publications to direct the readers to the corresponding dataset.

Most of the project (internal) operational data acquired and generated by the beneficiaries will be stored in the NextCloud online repository in a Cloud-Based System set up by the Project Coordinator and will be accessible anywhere from any device to anyone having assigned access rights.

The project website (<https://inplastwin.eu/>) provides important information about the project, and it will also be used to provide access to public deliverables and dissemination materials.

3.2.2. Data

Data related to the project dissemination activities, training material, open access publications, results, public deliverables, experiments description, and metadata will be accessible by default unless there is a specific reason not to publish the data. The open data will be deposited on Zenodo as a certified open-access repository where the data will be shared. The data will be accessible upon free and open access principles in Zenodo. The initial access to the data will be without any login or user account. The data used in scientific publications, posters, and oral communications will be made available for re-use as soon as is reasonably possible.

The level of data made available will also be considered; for example, pre-processed data will not be provided unless there is a clear reason for doing so. On request, raw data for practical reasons will be stored on the internal network drives owned by each beneficiary and can be accessed on request. In addition, IP-related data will be accessible in a restricted access repository, and the data access for research purposes during the project will be openly shared upon request by email.

Generally, data will be accessible via the common internet browser. Data will be published using standard file formats (txt, csv, etc.). All data will be accessed using standard tools that will be human and computationally readable. There are no restrictions on the use of the published data, but users will be required to acknowledge the project and the source of the data in any resulting publications, according to the CC-BY 4.0 license.

This consortium will not establish a data access committee. The coordinator, with support from the project partners, will have overall responsibility for the management of data/research outputs and quality assurance. The coordinator will be responsible for coordinating updates to the data management plan and for deciding on a case-by-case basis which data/research outputs will be kept and for how long. The participant(s) that produced the data will be responsible for organizing backup and storage, archiving, and depositing the data/research outputs within the chosen repositories.

3.2.3. Metadata

Metadata of deposited publications will be open under a Creative Common Public Domain Dedication (CC 0), in line with the FAIR principles and provide information at least about the following: publication (author(s), title, date of publication, publication venue); Horizon Europe funding; grant project name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action and, if possible, for their organisations and the grant.

Metadata will be guaranteed to remain available even after data is no longer available.

Documentation for any software required to access or view the data will be provided, preferably utilizing widely available software (either open source or commercial).

3.3. Making data interoperable

To maintain data impartiality and accuracy, InPlasTwin will employ triangulation between sources wherever possible and ensure result confirmation across different data/information sources using standardized file formats and vocabularies.

We will follow standard ontologies that describe experimental data and specify units to ensure interoperability and allow data exchange and reuse within and across disciplines. Data will be stored in commonly used formats, such as CSV and TXT files, making it accessible to various software. If the use of uncommon or project-specific ontologies or vocabularies is unavoidable, we will provide mappings to more widely recognized ontologies.

Our data will include qualified references to other datasets, including those from the current project or previous research, and will utilize standard vocabularies and machine-readable formats, with each dataset having a unique ID.

3.4. Increasing data re-use

The project aims to make data openly available for study, teaching, research purposes, and governance. Licensing or other means will be used to facilitate third-party access and reuse. The data will either be licensed under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a license with equivalent rights as set out in the Grant Agreement. Owners are suggested to adopt one of the Creative Commons 4.0 licenses, CC-BY-SA-4.0 most preferable, yet reasoned deviations may exist. This approach ensures the protection of knowledge and legitimate commercial interests while facilitating the free circulation of documents and preserving authors' and the project's intellectual property rights (CC BY-NC, CC BY-ND). Users will be required to acknowledge the consortium and the source of the data in any resulting publications. Alternatively, the Creative Commons Public Domain Dedication License (CC 0) will be used.

Moreover, we will take measures to safeguard the project's and Commission's image, as well as ensure the integrity of content, preventing de-contextualization. To promote the project's resources effectively, we will utilize the European Open Science Cloud (EOSC) Portal.

To enable data analysis and facilitate data re-use, short README files with information on methodology, units of measurements, analyses, variable conditions, etc., will be provided together with the data.

Any data published in open-access journals will be usable by third parties after the datasets have been deposited in the repository. The data that do not relate to peer-reviewed publications will be made available for re-use on a case-by-case basis.

The provenance and context of the data will be thoroughly documented to meet relevant standards using the Provenance and Context Content Standard (PCCS) Matrix. Data will be accompanied by information on how they were captured, processed, analysed, and validated. Other information that enables interpretation and use will also be provided.

Data quality will be assured through several quality assurance procedures:

- ✓ Repeated and comparison measurements,
- ✓ Adherence to standards for data recording,
- ✓ Use of controlled vocabularies and standard terminology,
- ✓ Validation of the data collected
- ✓ Provision of test results along with the data

- ✓ Peer-review of publications based on the data

The estimated costs for making the data and other research outputs FAIR are mostly related to personnel costs involved in collecting and managing data. These costs are included in the project's budget and will be claimed if compliant with the Grant Agreement's conditions. The participants will store other research outputs on their organisations' networks, which are protected by firewalls, backups, etc. The coordinator has also established a NextCloud online repository with a password-protected login for sharing, collaborative editing and storing the project-related data, accessible to all project partners. Where feasible, long-term preservation will be ensured by depositing the other research outputs in repositories. The coordinator will decide on a case-by-case basis on which other research outputs will be deposited and for how long.

The following issues could impact the sharing of other research outputs:

- ✓ Information relating to other research outputs acquired from third parties (e.g., manufacturers) will not be shared without their explicit consent.
- ✓ Information relating to other research outputs collected by the consortium at commercial sites will not be shared without the site owner's explicit consent.

The project will not share other research outputs with identifiable personal information. Sensitive information relating to the other research outputs will be collected, separated as soon as possible, and kept secure. Please also refer to the information provided in Section 7: Ethics.

4. Other research outputs

If needed, beneficiaries will also consider the management of other research outputs that will be generated or reused throughout the projects. Such outputs will be either digital (e.g. workflows, protocols, models, etc.) or physical (e.g. new materials, samples, videos, presentations, etc.).

The same FAIR principles will be applied to the digital research outputs as to the published research results and will be available for re-use.

5. Allocation of resources

The costs for making data and other research outputs FAIR will be kept to a minimum by using free repositories (ZENODO/OpenAIRE/Open Research Europe) and by making only relevant data and other outputs FAIR. Free and open-access repositories will be utilized whenever possible to minimize curation and storage costs.

Other costs include expenses for the personnel involved in data collection and management, as well as the costs associated with setting up and maintaining the project website and fees for publishing peer-reviewed scientific articles in full open access. Any costs for making research data and other project outputs FAIR are included in the project's budget and will be claimed if compliant with the Grant Agreement's conditions.

The project coordinator from JSI will have overall responsibility for the data management within the InPlasTwin project and will be responsible for coordinating updates to the data management plan. The coordinator will be responsible for organising data backup and storage, data archiving, and depositing the data within the repositories. Each beneficiary should respect the policies set out in this DMP. Datasets have to be created, managed, and stored appropriately and in line with the European Commission and local legislation. Dataset validation, registration of metadata, and backing up data for sharing through repositories are the responsibility of the partner that generates the data.

The datasets in Zenodo will be preserved in line with the European Commission Data Deposit Policy. The data will be preserved indefinitely (minimum of 5 years) and there are currently no costs for archiving data in this repository. Long-term preservation of other project-related data will be ensured by depositing the data within the coordinator repository. There are no costs associated with the long-term preservation of the data in these repositories. The coordinator will decide on a case by case basis on what data will be kept and for how long.

6. Data security

Storing and ensuring data security and integrity during research requires specific measures. Data should be easy to reach but at the same time safe and robust. Data should be protected from unauthorized access, changes, or loss.

Deposition of the data generated and scientific outputs of the project in public trustworthy repositories such as Zenodo will provide long-term preservation and curation in compliance with EU guidelines, as they have multiple replicas in a distributed file system that is backed up on a nightly basis.

In relation to personal data protection, all necessary actions will be taken within the project management and by all beneficiaries to ensure compliance with applicable European and national regulations and professional codes of conduct. The participants will store and share data on their organisations' in-house networks, which are protected by firewalls, backups, etc., following well-defined internal data preservation. Access is protected by username & password, as well as user rights to the project drive. The coordinator has also established a NextCloud online repository with a password-protected login for sharing, collaborative editing and storing the project-related data, accessible to all project partners. This repository is provided, administrated, hosted, and maintained by the IT department of JSI in compliance with GDPR.

InPlasTwin recognizes that the data provided by individuals and organizations remains their property. As such, the project will handle the data under the responsibility of the InPlasTwin Consortium. To formalize these arrangements, a Consortium Agreement was signed by all project partners before the commencement of the project, outlining the rights and responsibilities regarding data ownership, access, and usage.

7. Ethics

To ensure the protection of the personal data of participants involved in the educational and training activities, workshops, summer school, and other relevant activities, InPlasTwin will cater to:

a) **Obtaining informed consent:** Before collecting any data, the project will make sure to obtain informed consent from the participants. This means explaining the purpose of the research, what data will be collected, and how they will be used. Participants will also be informed of their right to withdraw from the project activity at any time. Informed consent will be collected through written or electronic consent forms.

b) **Data anonymization:** The project will remove any identifying information from the data, such as names, addresses, and phone numbers. Pseudo-anonymization will be achieved by assigning unique identifiers to participants that are not linked to their personal information.

c) **Secure data storage:** Collected data will be securely stored. This means using password-protected computers and storage devices, as well as encrypting any electronic data collected from participants.

d) **Minimizing collected data as well as minimizing access to data:** Only data that are relevant to InPlasTwin will be collected (i.e., related to the research and education practices in widening and non-widening countries regarding MNPs). Also, only persons working on the project should have access to the data collected. Before sharing data, it shall be ensured that collaborators also adhere to data protection guidelines.

e) **Appropriate data disposal:** Once the project is complete, anonymized data will be stored as a reference for future research unless participants have opted in for their retention.

Management of the personal data will comply with the European laws about data security and the protection of privacy as well as institutional national guidelines, implementing the General Data Protection Regulation (GDPR).