**Funded by
the European Union**

D5.1 DATA
MANAGEMENT PLAN

IPPT PAN

Author list

Name	Organization
Marika Kowalska	IPPT PAN
Dorota Omelczuk	IPPT PAN
Marta Jurkowska	IPPT PAN

Document History			
Version	**Date**	**Note**	**Revised by**
01	24.06.2025	First Draft	IPPT PAN
02	27.06.2025	Final Version	IPPT PAN

Disclaimer

The content of the publication herein is the sole responsibility of the publishers, and it does not necessarily represent the views expressed by the European Commission or its services.

While widerAdvance Facility is 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 the European Research Executive Agency (REA). Neither the European Union nor the European Research Executive Agency (REA) can be held responsible for them.

While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the widerAdvance Facility consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the widerAdvance Facility Consortium nor any of its members, their officers, employees or agents shall be responsible for or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the widerAdvance Facility Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Executive Summary

This Data Management Plan outlines the strategy of the widerAdvance Facility for handling data and other research outputs in accordance with Horizon Europe's FAIR principles and open science requirements. Although the project does not generate scientific datasets, it will produce a wide range of operational and strategic data as well as digital tools and training materials.

The DMP details the procedures for data collection, storage, accessibility, licensing, metadata standards, ethical compliance, and long-term preservation, ensuring that all outputs are managed responsibly and contribute to the sustainability and visibility of research valorisation in Widening and Outermost regions.

Table of Contents

List of Acronyms

Acronym	Full Term
CSA	Coordination and Support Action
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable, Re-usable
GDPR	General Data Protection Regulation
HE	Horizon Europe
H2020	Horizon 2020
IP	Intellectual Property
JSON	JavaScript Object Notation
KER	Key Exploitable Result
NDA	Non-Disclosure Agreement
DOI	Digital Object Identifier
WP	Work Package
ESR	Evaluation Summary Report
CC0	Creative Commons Zero – Public Domain Dedication
CC-BY	Creative Commons Attribution License
CC-BY-NC-SA	Creative Commons Attribution-NonCommercial-ShareAlike License
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
ERA	European Research Area
R&I	Research and Innovation
EC	European Commission
EU	European Union
CSV	Comma-Separated Values
XLSX	Microsoft Excel Open XML Spreadsheet
DOCX	Microsoft Word Open XML Document
PDF	Portable Document Format
MP4	MPEG-4 Video File
XML	Extensible Markup Language

1. Data summary

The widerAdvance Facility project will generate a wide spectrum of data, both qualitative and quantitative, to support its primary mission of enhancing the visibility, valorisation, and replication of results from Horizon 2020 and Horizon Europe Widening actions. As a Coordination and Support Action (CSA), the project is not expected to produce original research datasets in the scientific sense. Nevertheless, it will generate substantial operational, strategic, and stakeholder-related data.

1.1. Re-use of existing data

The project will not re-use experimental research data from previous scientific studies. However, it will make strategic use of publicly available contextual information and insights from previous initiatives such as IP Booster, Horizon Results Booster, and selected Widening projects (especially those supported under H2020 and Horizon Europe). These inputs will be analysed to identify best practices and lessons learned, which will inform the design of tools and services offered by the Facility, including capacity-building materials, methodological templates, and evaluation frameworks. The reuse of such materials is justified by the Facility's objective to scale up proven practices and tools in a harmonised way across Widening countries.

Where the re-use of existing datasets has been considered but ultimately discarded, this decision has typically been based on two factors: either (a) the data is no longer aligned with current policy frameworks or target groups; or (b) the data structure and licensing conditions do not support meaningful interoperability and reuse within the Facility's technical environment.

1.2. Types and formats of data

The types of data expected to be generated or processed in the project include:

- • Administrative and coordination data: meeting agendas, minutes, and reports in standard document formats (e.g. DOCX, PDF);
- • Stakeholder data: contact and engagement records stored in spreadsheet or structured text formats (e.g. XLSX, CSV);
- • Evaluation data: forms, surveys, and summary tables supporting feedback and impact assessment;
- • Dissemination content: newsletters, presentations, toolkits, and visuals shared in commonly used digital formats;
- • Multimedia: recordings of events and interviews provided as video files or hosted online;

• Metadata: basic descriptive information to support data discovery, including titles, authorship, date, and licensing.All data will be organised in compliance with FAIR principles and GDPR obligations.

1.3. Purpose and relation to project objectives

The generation, processing, and analysis of data within the widerAdvance Facility is intrinsically linked to the achievement of the project's five strategic objectives, each addressing a critical need in the valorisation and dissemination ecosystem of Widening and Outermost regions. The data strategy not only underpins operational execution but also ensures the long-term sustainability, visibility, and policy relevance of project outcomes.

In support of Objective 1 (O1) - to improve the exploitation of Key Exploitable Results (KERs) - data will be gathered through intake forms, evaluation frameworks, service tracking sheets, and feedback mechanisms associated with the Academy, Pitching, IP, and standardisation services. This includes documenting participation rates, KER characteristics, support intensity, and service outcomes. The structured data will enable the monitoring of how individual KERs are supported across services and allow for portfolio-level analysis of their technological maturity, societal value, and market readiness. It also facilitates iterative improvement of support protocols and benchmarking of best practices across regions.

For Objective 2 (O2) - to ensure the long-term effect of improved exploitation and dissemination through education - the project will collect data related to educational and training interventions, including participant profiles, training attendance, learning assessments, and post-training follow-up surveys. These datasets will help assess learning outcomes, adjust future content, and validate the long-term effectiveness of structured learning formats such as webinars, workshops, and study visits.

The achievement of Objective 3 (O3) - to improve visibility and recognition of Widening and Outermost regions - will rely on data from storytelling initiatives, regional case studies, synergy award documentation, and media analytics. This includes qualitative and quantitative data collected through stakeholder interviews, success story submissions, social media metrics, and event reports.

In the context of Objective 4 (O4) - to provide beneficiaries with tools and opportunities for sustainable dissemination and exploitation - data management will focus on monitoring matchmaking event participation, feedback on tool usability, and engagement metrics for the synergies matrix and coaching sessions.

To fulfil Objective 5 (O5) - to deliver input to EU, national, and regional valorisation policy - the project will collect and analyse data related to barriers to dissemination and exploitation, stakeholder needs, and policy feedback mechanisms. This includes outputs from structured surveys, thematic consultations, and targeted interviews.

In all cases, the data collected will be handled in compliance with GDPR and FAIR principles, ensuring both legal conformity and maximum utility.

1.4. Expected size of the data

While the exact data volume will depend on stakeholder engagement levels and the scale of dissemination and evaluation activities, the estimated volumes include:

- Text-based documents: ~100 MB
- Spreadsheets and datasets: ~1 GB
- Multimedia files (recordings, animations): 5–10 GB
- Metadata: small in size but structurally essential

These figures are projections and will be updated during periodic reviews of the DMP.

1.5. Origin of the data

Data will originate from:

- Consortium members, through their implementation of project tasks and internal reporting;
- External stakeholders, via registration forms, consultation feedback, matchmaking activities, and submission of KERs for review;
- Publicly available sources, such as previous Horizon Europe projects, existing repositories and datasets made available by the European Commission or national agencies.

1.6. Data utility beyond the project

The data produced by widerAdvance Facility will be of strategic importance and long-term utility for multiple audiences beyond the project consortium. These include:

- Research managers and administrators in Widening countries;
- Policy makers at EU, national, and regional levels;
- Innovation support organisations;
- Future CSA or R&I projects;
- Researchers analysing research impact or valorisation processes.

All data made publicly available will be accompanied by appropriate metadata, licensing information, and persistent identifiers to ensure discoverability and reuse.

2. FAIR data

2.1. Making data findable, including provisions for metadata

The widerAdvance Facility is committed to making all public project data and outputs findable and accessible through the use of persistent identifiers, rich metadata, and adherence to recognised standards. These practices are essential to ensuring long-term discoverability, traceability, and reusability of the knowledge and resources generated throughout the project's lifecycle.

All data shared publicly will be registered with a Persistent Identifier (DOI), assigned via trusted repositories like Zenodo. Metadata will follow recognised open standards to ensure traceability and reusability.

Metadata will include:

- Title, creator(s), project number and acronym,
- Versioning and licensing information (CC-BY, CC0, or CC-BY-NC-SA),
- Keywords (e.g., "Widening", "valorisation", "stakeholder engagement"),
- Links to related resources (e.g., ESRs, success stories, evaluations).

All metadata will be indexed in open repositories to support discoverability. Specific templates will be described where necessary to support clarity and reuse.

A naming convention based on WP/task codes, partner abbreviations, document type, and versioning will also be maintained as defined in the project's Quality Assurance & Risk Management Plan (D5.2).

2.2. Making data accessible

2.2.1. Repository

In accordance with Horizon Europe's open science policy and the FAIR principles, the widerAdvance Facility is committed to ensuring that its outputs and datasets are openly accessible through reliable, long-term, and standards-compliant repositories. The project will adopt a dual dissemination strategy that combines trusted external repositories with a dedicated project portal, thereby maximising discoverability while meeting legal and ethical obligations.

The primary repository for public data and documentation will be Zenodo, an open-access platform developed under the OpenAIRE programme. Zenodo complies fully with FAIR data standards and

assigns a Digital Object Identifier (DOI) to each uploaded record, ensuring that resources are persistently accessible, properly cited, and indexed through their own dedicated landing pages. Its affiliation with the European Commission and commitment to long-term preservation make it especially well-suited for archiving widerAdvance outputs.

Complementing Zenodo, the widerAdvance Facility portal will function as a curated access point for project deliverables, training materials, multimedia content (e.g. webinar recordings), and other knowledge assets. It will offer user-oriented navigation, contextual descriptions, and direct hyperlinks to datasets archived in Zenodo, creating a seamless interface between internal infrastructure and open science platforms.

Repository arrangements were planned early in the project's lifecycle to ensure technical compatibility, ease of access, and sustainability. Zenodo's features also support access restrictions where required. In such cases - such as datasets involving personal data, stakeholder submissions, or intellectual property-sensitive content - the project will apply embargoes or closed access settings, with conditions transparently documented in metadata.

Non-public data will be stored securely on institutional servers maintained by IPPT PAN (Data Manager) and within the password-protected backend of the Facility portal. Access will be role-based and granted only under GDPR-compliant procedures and with documented legitimate interest.

This hybrid data dissemination model - integrating high-trust public repositories with a structured, managed portal - ensures both openness and security. It enables broad accessibility to project results while safeguarding sensitive information, thereby fulfilling the widerAdvance Facility's commitment to responsible, FAIR-aligned, and ethically sound data stewardship.

2.2.2. Data

The widerAdvance Facility adheres to the Horizon Europe principle of making data "as open as possible, as closed as necessary." The majority of data generated during the project - particularly materials such as training toolkits, communication templates, policy briefs, evaluation frameworks, and public deliverables - will be made openly available through mentioned repositories, and the project's own widerAdvance.eu portal.

However, there are important categories of data that will be subject to restricted access conditions due to:

1. **Legal requirements (GDPR compliance)**:
   Personal data (e.g. names, institutional affiliations, and email addresses of stakeholders, trainees, applicants, or workshop participants) will be collected and processed in accordance with

the General Data Protection Regulation (GDPR). This data will not be made openly available, and access will be limited to project staff with a legitimate role in its processing. Anonymised or aggregated statistics may be published when feasible.

2. **Contractual obligations and IP sensitivity**:
   Some datasets, particularly those related to Key Exploitable Results (KERs) provided by stakeholders during individual support sessions, will be considered confidential and protected under Non-Disclosure Agreements (NDAs). These materials may include unpublished findings, innovation descriptions, or commercialisation strategies. The Consortium Agreement permits such data to remain closed to preserve the legitimate interests of the contributor.

3. **Strategic project considerations**:
   Certain documents and intermediate outputs (e.g. internal analysis grids, draft policy briefs, unpublished versions of impact reports) may be held under embargo until final validation is completed or accompanying publications are released. Embargoes may also be applied temporarily to allow contributors time to pursue IP protection, though patenting is not a primary activity in this CSA.

   When applied, embargo periods will be time-bound and justified in the DMP updates, typically lasting no more than 6–12 months post-creation or until formal publication.

All openly accessible data will be made available via a free and standardised protocol, such as HTTPS, through Zenodo or the project portal. These platforms comply with EU open data standards and ensure persistent identifiers (DOIs) and machine-readable metadata.

For restricted data:

- Controlled access will be implemented via platform and IPPT PAN's secure servers;
- Access to sensitive materials will be role-based, and granted only after approval by the Data Manager or designated WP Lead;
- Identity verification will rely on authenticated institutional logins, access logs, and documentation of legitimate interest (e.g. participation in a related work package or signed NDA).

At this stage, there is no formal Data Access Committee foreseen, given the project's scope and the limited volume of sensitive data. However, in cases where sensitive personal or commercially valuable data is requested, the Data Manager – Ms. Dorota Omelczuk (IPPT PAN), in consultation with the relevant WP leader, will evaluate requests on a case-by-case basis to determine if and under what conditions data access can be granted.

All procedures will be documented and updated in future versions of the DMP to reflect evolving data types and stakeholder requirements.

2.2.3. Metadata

In line with Horizon Europe's open science policy and the widerAdvance Facility's commitment to the FAIR principles, all metadata associated with publicly shared datasets and research outputs will be made openly available and licensed under the Creative Commons Public Domain Dedication (CC0), unless legal or contractual constraints apply. This approach ensures metadata can be freely harvested, reused, indexed, and cited without restriction, even in cases where the underlying datasets are restricted or embargoed.

Each metadata record will follow open and widely accepted metadata standards to support discoverability, and will include basic descriptive elements such as title, author, license, and version. Specifically, metadata will provide:

- A clear title and abstract;
- Names and affiliations of the creators and contributors;
- The project name, acronym, and grant number (101159516);
- Date of creation and version;
- Subject classification and keywords;
- Licensing terms;
- A persistent identifier (e.g. DOI);
- A direct access link to the dataset, or in the case of restricted access, clear instructions for requesting access.

Metadata will also describe any technical requirements needed to access or interpret the associated data, including file formats, software environments, or dependencies. Where applicable, references to standard open formats where applicable, will be provided. These may either be included directly in the metadata record or linked to external repositories. Where relevant, the required software or format will be clearly indicated.

Importantly, metadata will remain persistently available, even if the associated datasets are withdrawn (e.g. due to data retention limits or embargo expiry). Long-term preservation will be ensured through trusted repositories, which are committed to maintaining metadata will be retained to support long-term access and traceability. This guarantees traceability, and a documented history of research outputs, contributing to long-term accountability and transparency.

In cases where datasets cannot be made openly available - such as those involving personal data, confidential stakeholder input, or IP-sensitive content - the metadata will still be published, clearly indicating the restricted access status and providing relevant contact details for access requests. This ensures transparency while upholding data protection obligations and the legitimate interests of contributors.

By ensuring that all metadata is openly licensed, persistently hosted, and richly documented, the widerAdvance Facility will support not only compliance with legal and policy frameworks, but also promote

broad data discoverability, interoperability, and responsible reuse across the European Research and Innovation community.

2.3. Making data interoperable

The project uses commonly accepted file formats (e.g. CSV, XLSX, PDF, MP4) to support compatibility and reuse across platforms. All shared data will be structured in a way that enables easy access and understanding by others. If project-specific templates or data structures are developed, they will be clearly described to ensure clarity and reusability. Any future updates will be reflected in later versions of the DMP.

2.4. Increase data re-use

The widerAdvance Facility is committed to maximising the reuse potential of all datasets and knowledge assets produced throughout the project. To enable responsible and effective reuse, all data made publicly available will be appropriately licensed, accompanied by rich documentation, and structured to ensure long-term usability by third parties, in line with FAIR principles and the Horizon Europe Grant Agreement.

Openly shared data will be published under standard licences, such as:

- **CC-BY** for resources requiring attribution;
- **CC0** for metadata and other content intended for unrestricted reuse;
- **CC-BY-NC-SA** for outputs meant for non-commercial replication, such as stakeholder toolkits or training materials.

Data that is subject to legal, ethical, or contractual restrictions - such as personal data protected under GDPR or confidential stakeholder inputs (e.g. from KER support sessions or IP-sensitive services) - will not be made openly accessible. Such datasets will be stored securely and retained for internal use only. Metadata records will still be published in trusted repositories, clearly indicating the access restrictions and contact information for justified access requests.

To facilitate reuse and reproducibility, every dataset published will be supported by:

- A detailed readme file providing context, methodology, and file structure;
- Definitions of variables and units of measurement for structured datasets;
- Where applicable, codebooks and explanations of data classifications or scoring systems;
- Information on data processing steps, including cleaning or transformation;
- Specification of any required software or file format needed to access or analyse the data, with open alternatives suggested where feasible.

The provenance of all data will be fully traceable through metadata, which will record the origin, creators, institutional affiliations, funding references (including GA number 101159516), and the relevant work package or task. Derived datasets will include references to their original sources to ensure transparency.

Data quality will be assured through internal reviews led by work package teams and coordinated by the Data Manager. This includes peer validation of key outputs, systematic checks on dataset accuracy, and regular reviews to ensure coherence with the project's data management standards. Task leaders will report any quality concerns, and corrective actions will be coordinated centrally.

Where necessary, embargoes may be applied to selected datasets to allow for validation or protect confidentiality. These embargoes will be clearly indicated and limited in duration, with updates provided in subsequent DMP versions. Embargo status and expiration will be tracked via the project's internal documentation register and reflected in metadata updates on Zenodo

All data intended for reuse will remain available for at least 5 years after the project's conclusion, in certified repositories and through the widerAdvance portal, ensuring their continued relevance for future EU-funded projects, policy development, and capacity-building initiatives in the Widening community.

3. Other research inputs

Although the widerAdvance Facility is not expected to generate traditional research datasets, it will produce a wide range of non-data research outputs, many of which will hold significant strategic value for replication, capacity building, and policy uptake. These outputs span digital tools, methodologies, training assets, and stakeholder engagement models - all of which fall under the broader umbrella of research outputs subject to FAIR-aligned management practices.

The project will ensure that these outputs are created, documented, and disseminated in ways that promote accessibility, reusability, and transparency, in accordance with Horizon Europe guidelines.

3.1. Reuse and sharing strategy

All research outputs will be made available through the widerAdvance portal, supported by metadata, user guidance, and contextual information. Where public hosting is not feasible (e.g. for tailored, sensitive materials), access will be mediated through request protocols or shared via controlled-access repositories.

Post-project, the sustainability strategy includes keeping these outputs accessible for at least five years, with hosting provided by institutional repositories and Zenodo. Reuse will be actively encouraged, including:

- Adaptation by future Horizon Europe projects;

- Uptake by national or regional R&I support schemes;
- Integration into institutional training or policy development efforts.

Through this comprehensive approach, widerAdvance ensures that all its research outputs - not just data - contribute to long-term systemic change and capacity building across Europe's widening R&I landscape.

4. Allocation of resources

The widerAdvance DMP implementation will be supported through:

- Partner-level personnel (Data Manager, WP Leaders, IT support),
- Use of free repositories (e.g., Zenodo, GitHub),
- Internal server infrastructure (IPPT PAN, META Group),
- Readme creation, metadata population, GDPR documentation.

No dedicated budget is assigned; costs are absorbed within WP5 and respective task budgets. The plan is compliant with the Grant Agreement and does not require separate financial allocation.

Person responsible for data management in widerAdvance Facility project is Ms. Dorota Omelczuk (IPPT PAN).

5. Data security

Data related to the GDPR will be archived and locked with a password. The password will be known to the data processor and the Coordinator. All files will be under the management of access rights by the Consortium Member on whose servers these files are stored.

Data Manager periodically conducts data security checks.

Each Consortium Member hosting data is responsible for maintaining its ongoing security and ensuring access is restricted appropriately. Each data record will be subject to regular access audits as part of standard security checks.

6. Ethics

The widerAdvance Facility is committed to upholding the highest ethical standards in all aspects of its data collection, processing, sharing, and preservation activities. While the project does not engage in high-risk research involving vulnerable populations, biomedical data, or experimental human subject research,

ethical and legal considerations still apply, particularly with regard to the management of personal data collected during stakeholder engagement, training activities, surveys, and matchmaking services.

6.1. Ethical and legal considerations for data sharing

The project does not foresee major ethical risks related to data management; however, it acknowledges the need to handle certain categories of data - particularly names, affiliations, roles, and email addresses of participants - in full compliance with GDPR (Regulation EU 2016/679) and national data protection regulations applicable in partner countries.

Although the widerAdvance project was not required to include a dedicated Work Package on ethics, and no formal ethics deliverables were mandated in the Grant Agreement, ethical safeguards are embedded into the project's operational framework. These are also referenced in Section 1.2 and Task 5.5 of the Description of Action (DoA), where the importance of secure, confidential handling of stakeholder data is emphasised, especially in the context of knowledge transfer and service delivery.

7. Other issues

In addition to compliance with Horizon Europe data management requirements and FAIR data principles, the widerAdvance Facility will adhere to a set **of** national, institutional, and sector-specific procedures that further strengthen the quality, consistency, and legal compliance of its data governance practices.

Follow us on:

@widerAdvance

wider Advance Facility

@wideradvance.bsky.social

Learn more:

www.widerAdvance.eu