



WIDE AcrossEU



Horizon Europe Data Management Plan

25 October 2024

History of changes

HISTORY OF CHANGES		
<u>VERSION</u>	<u>PUBLICATION DATE</u>	<u>CHANGE</u>
1.0	11.09.2024	Initial version.
1.1	25.10.2024	The version uploaded as a deliverable (WP6: D6.1).

Contributors

The following contributors are related to the project of this DMP:

- Monika Vejchodova
monika.vejchodova@upce.cz, ORCID: [0009-0007-3870-2365](https://orcid.org/0009-0007-3870-2365)
Roles: Contact Person, Data Curator, Data Manager, Project Leader
Affiliation:

University of Pardubice (UPCE)

- Diana Kalášková
diana.kalaskova@upce.cz, ORCID: [0000-0002-1550-731X](https://orcid.org/0000-0002-1550-731X)
Roles: Contact Person, Data Manager
Affiliation:

University of Pardubice (UPCE)

Projects

We will be working on the following projects and for those are the data and work described in this DMP.

WIDEn performance in research an innovation capacity and competence Across EU

Acronym

WIDE AcrossEU

Start date

2024-05-01

End date

2027-08-31

Funding

Research Executive Agency (European Union)

Project number

101158561 (granted)

Project Coordinator

UNIVERSITY OF PARDUBICE, (abbr. UPCE)

For the DMP creation and update represented by Ing. Monika Vejchodová

Project Partners

- Ss. CYRIL AND METHODIUS UNIVERSITY IN SKOPJE (abbr. UKIM)
For the update of DMP represented by Beti Kostadinovska Dimitrovska
- ZHYTOMYR POLYTECHNIC STATE UNIVERSITY (abbr. ZPSU)
For the update of DMP represented by Yuliya Bogoyavlenska
- LAPIN YLIOPISTO (abbr. ULAPLAND)
For the update of DMP represented by Hari Malinen represented by Marlene Kohllechner-Autto
- UMEA UNIVERSITET (abbr. UMU)
For the update of DMP represented by Agneta Hanell Plamboeck

Project Summary

Our project represents the opportunity to strengthen institutions from Widening countries and upscale them on higher level through capacity, skills and competence building and networking actions. The consortium consisting of partner institutions from Czechia, North Macedonia, Ukraine, Finland and Sweden have diverse experience with participation in Horizon Europe (HE), Synergic EU R&I funding, Smart Specialisation Clusters and other international R&I funding schemes. The key objective of the project “WIDE AcrossEU” is to move the participants from Widening countries (the core group - 3 higher education institutions from the Czech Republic, North Macedonia and Ukraine), who are formerly a current single beneficiaries of regional funding programmes (ERDF, IPA, Interreg and similar investments for R&I infrastructure), to higher level and prepare them for wider and more successful participation in direct EU funding programmes.

The specific objectives, which are: 1. To map the needs and requirements of target groups in Widening partners, 2. To strengthen competitiveness of target groups in Widening countries by acquiring new transferable skills for R&I staff, 3. Access to networks and partnerships to secure more successful proposal and create a base for lead partner roles for Widening partners, 4. To better use R&I infrastructure funded under ERDF or similar investments, 5. To create strategies for R&I and Human Resources, 6. To improve access to excellent European R&I networks and communities. These will be achieved through the following work packages: WP1: Need of analysis for capacity and competence building, WP2: Acquisition of new transferable skills, WP3: Exploitation of R&I funding opportunities & networks, WP4: R&I valorisation, policy support and synergies, WP5: Dissemination, communication and exploitation, WP6: Management and coordination.

The project will last 40 months.

1. Data Summary

Non-equipment datasets

We also collect data from questionnaires and case report forms. The non-equipment datasets are:

- **Questionnaires** – Right now, there is no data which could be used as a resource for the activities planned in the project. That is why, during the project implementation phase, we will collect data from questionnaires to be able to prepare training programme for researchers and supporting staff. The primary purpose of data collection in this project is to support the development and implementation of a comprehensive training program tailored to the needs of researchers and supporting staff. This is directly related to the project's objective of enhancing research capabilities through targeted training and support. The research activities of this project include carrying out two questionnaires (WP1): for the analysis of the needs, requirements and expectations of target groups. The data collected from the questionnaires will provide insights into the needs, requirements, and expectations of the target groups, forming the foundation for developing customized training programs and other support mechanisms: information about grant programmes, calls, portals, project preparation, project writing, project implementation, project management, project finishing, reporting etc. This data will be completely anonymised and no tracing back to the people who completed the questionnaire will be possible.
- **The reports on training and short-term stays** (WP2) will be uploaded in a form of a document. This document will be based on anonymized training evaluation reports and evaluation reports of short-term stays of participants.
- An **online R&I infrastructure online platform** will be created and maintained (WP5). This platform is meant to facilitate future collaborations and make finding possible project partners easier. The data in stores will include equipment, installation, database, knowledge base, virtual resource or other resource, and the associated human competence necessary to carry out research of the highest quality. However, no personal data will be stored in this database as it will contain the links to the relevant workplaces/departments.

Data formats and types

The project will generate textual and numerical data through questionnaires, training reports, and evaluation reports. The data formats will include:

- docs, pdf, rtf, txt, xlsx
It is a standardized format. This is a suitable format for long-term archiving. We expect to have 10 GB of data in this format.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

All data produced during the project that does not contain personal or sensitive information will be made openly available. This includes:

1. **Evaluation Data from Questionnaires:** insights from researchers and supporting staff on their interest in education, internships, shadowing, mobilities, and their prior experiences with grant programs and project management.
2. **Information on Grant Programs and Calls:** information regarding available grant opportunities and application processes.
3. **Information on Relevant Events and Workshops:** details on upcoming workshops, educational courses, and events relevant to the research community. There are no legal, contractual, or voluntary restrictions that prevent the sharing of these datasets, as they do not contain any sensitive or personal data. If, during the project, certain data is identified as restricted due to unforeseen legal or ethical reasons, access will be limited accordingly, and these conditions will be clearly specified. (published)

The dataset has the following identifiers:

- DOI: <https://doi.org/>

We will distribute the dataset using:

- *Domain-specific repository:*
- **Zenodo** (Zenodo)

There is no need to contact the repository beforehand. A persistent identifier will be assigned by the repository as Zenodo assigns DOI upon upload. The repository will make sure that the persistent identifier can be resolved to a digital object.

The assigned persistent identifier is specified: 10.5281/zenodo...

There will be different versions of this data over time; the versions will be dated. We will not be adding a reference to any data catalogue because the data will be stored in a repository that is the prime source of data for re-use in the field.

There are the following 'Minimal Metadata About ...' (MIA...) standards for our experiments:

- **Dublin Core Metadata Initiative Terms** (DCMI Metadata Terms)

We made a SOP (Standard Operating Procedure) for file naming. We will adhere to a structured naming convention for all files, ensuring consistency and clarity. The naming convention will be as follows: `yymmdd_MeaningfulShortName_ProjectName_Language` this format includes the date, a concise description, project acronym, and language code.

- Example: 241129_Questionnaire1Results_WIDEAcrossEU_EN.

We will be keeping the relationships between data clear in the file names. All the metadata in the file names also will be available in the proper metadata.

Due to the naming conventions following the pattern stated above (starting with the YYMMDD format), the chronological order of the individual versions of the data will be clear. This will help in tracking the evolution of the datasets over time and ensure that users can identify the latest version of the data.

2.2. Making data accessible

We will be working with the philosophy *as open as possible* for our data.

All of our data can become completely open over time.

Limited embargo will not be used as all data will be opened immediately.

Metadata will be openly available including instructions how to get access to the data. Metadata will be available in a form that can be harvested and indexed (managed by the used repository / repositories).

All data will be owned by the Principal Investigator.

The distributions will be accessible through:

- *Domain-specific repository:*

Zenodo (Zenodo)

The Project Coordinator will make use of its own digital library and also of the trusted repository, Zenodo.

We do not need to contact the repository because it is a routine for us. The distribution will be available under the following license:

- Freely available with obligation to quote the source (e.g. CC-BY). This license allows that as soon as it is uploaded to the project webpage and the Zenodo repository, it will allow for the broadest re-use, enabling others to distribute, remix, adapt, and build upon the data, as long as proper credit is given to the original creators. This approach ensures that the data can be widely used and shared by third parties in various contexts, fostering greater collaboration and innovation.
- We do not anticipate the need for an embargo period since there are no plans for patent applications or other restrictions that would delay data sharing. If an embargo is needed in the future, for example, to allow time for publishing related research articles, it will be clearly communicated, specifying the reason and duration.

Any user of this data can use it without any specific software.

The dataset will be published as soon as possible after collecting it.

When it comes to communication amongst the project partners and the coordinator, all the data will be shared within a MS TEAMS group limiting the access to only those who are part of the project consortium.

2.3. Making data interoperable

The project aims to ensure that the data and metadata produced is interoperable, facilitating data exchange and re-use among researchers, institutions, and organizations across different countries. To achieve this, we will be using the following data formats and types:

- docs, pdf, rtf, txt, xlsx
In a standardized format. And this format compatible with open software applications and commonly used in research settings. This will enable the combination and integration of our data with other datasets from different origins.

2.4. Increase data re-use

The metadata for our produced data will be kept as follows:

- **All data produced during the project that does not contain personal or sensitive information will be made openly available. This includes: (i) Evaluation Data from Questionnaires, (ii) Information on Grant Programs and Calls, (iii) Information on Relevant Events and Workshops**
- There are no legal, contractual, or voluntary restrictions that prevent the sharing of these datasets, as they do not contain any sensitive or personal data. If, during the project, certain data is identified as restricted due to unforeseen legal or ethical reasons, access will be limited accordingly, and these conditions will be clearly specified.
- This data set will be kept available as long as technically possible.
- The metadata will be available even when the data no longer exists.

For open-access data, no identity verification will be necessary. Since all shared data will be open access and free from restrictions, no data access committee is required. If any unforeseen legal or ethical concerns arise during the project, we will assess the restrictions required and ensure that these are communicated clearly to all stakeholders, with data access limited as necessary. Should the need to restrict access to certain datasets arise in the future, the corresponding steps will be taken, user identity verification will be handled through a data access request process, where users will need to provide their institutional affiliation and purpose of use. The requests for data will be evaluated on a case-by-case basis. This process will be managed by the project coordinator.

As stated already in Section 2.2, all of our data can become completely open over time.

We will be archiving data (using so-called *cold storage*) for long term preservation already during the project. The data are expected to be still understandable and reusable after a long time.

To validate the integrity of the results, the following will be done:

- We will run a subset of our jobs several times across the different compute infrastructures.
- We will be instrumenting the tools into pipelines and workflows using automated tools.
- We will use independently developed duplicate tools or workflows for critical steps to reduce or eliminate human errors.
- We will run part of the data set repeatedly to catch unexpected changes in results.

Metadata Methodologies:

All datasets will be saved in interoperable formats such as CSV and XLSX for numerical data and DOCX or PDF for textual data. These formats are widely supported by various software tools, including open-source options. Standard vocabularies and controlled terms will be used for metadata fields such as subject, author, date, and description.

- We will follow the Dublin Core metadata standard for describing datasets. This standard is widely used and supports interoperability across different disciplines and data repositories.

3. Other research outputs

We use Data Stewardship Wizard for planning our data management and creating this DMP.

4. 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.

- **WIDEn performance in research an innovation capacity and competence Across EU - 101158561**

In our project, the costs associated with making data FAIR (Findable, Accessible, Interoperable, and Re-usable) are minimal. The use of the digital library services of University of Pardubice and Zenodo repository, provide free access to its services, eliminate expenses related to data deposition and dissemination. Consequently, we do not anticipate any direct costs for data publication or storage.

Personnel costs related to data curation and DMP updates are considered part of the existing project staff roles and responsibilities, and therefore covered under the project's overall personnel budget.

If any unforeseen costs related to ensuring data FAIRness occur (e.g., specialized data curation tools, additional storage requirements), these will be covered under the project's budget allocated for data management activities. As stipulated in the Grant Agreement, costs related to open access to research data are eligible for funding, and we will ensure compliance with these conditions.

We will be archiving data (using so-called 'cold storage') for long term preservation after the project but also already during the project. The costs for archiving data will be paid out of departmental budgets from one or more of the project participants. The minimum lifetime of the archive is 10 years. The archival period can be extended – one of the principal investigators involved in the project will decide. The decision whether or not to extend the renewal be based on the actual use of the archived data. Data formats of data in cold storage will be upgraded if they become obsolete. Archived data will be migrated regularly to more modern storage media (e.g. newer tapes).

None of the used repositories charge for their services.

Monika Vejchodova is responsible for reviewing, enhancing, cleaning, or standardizing metadata and the associated data submitted for storage, use and maintenance within a data centre or repository.

Monika Vejchodova and Diana Kalaskova are responsible for maintaining the finished resource. They will provide relevant updates and data for the DMP as needed.

To execute the DMP, no additional specialist expertise is required. We do not require any hardware or software in addition to what is usually available in the institute.

5. Data security

Project members will not carry data with them (e.g. on laptops, USB sticks, or other external media). All data centres where project data is stored carry sufficient certifications. All project web services are addressed via secure HTTP (<https://...>). Project members have been instructed about both generic and specific risks to the project.

The risk of information loss in the project or organization is acceptably low. The possible impact to the project or organization if information is leaked is small. The possible impact to the project or organization if information is vandalised is small.

All personal data will be collected anonymously.

The archive will be stored in a remote location to protect the data against disasters. The archive needs to be protected against loss or theft. It is clear who has physical access to the archives.

The project has established comprehensive measures to ensure data security throughout the data lifecycle, including data recovery, secure storage, and transfer of sensitive data. Key provisions include:

- **Regular Backups:** All project data will be regularly backed up using automated systems at the project coordinating institution. Backups will be stored in secure, separate locations to ensure data can be recovered in case of accidental loss or system failure.
- **Certified Repositories:** Data will be stored in certified repositories, such as the Zenodo repository, which adheres to OpenAIRE guidelines and offers long-term preservation and secure access to research data.
- **Institutional Servers:** Sensitive data and intermediate datasets will be securely stored on the institutional servers of the project coordinating institution. These servers comply with institutional and European data protection regulations, ensuring the safe storage of data.
- **Encryption:** For sensitive data, encryption methods will be employed both during storage and transfer to prevent unauthorized access.
- **Data Transfer Protocols:** Secure file transfer protocols (e.g., SFTP, HTTPS) will be used for transferring data between project partners and external stakeholders.
- **Access Control:** In case of generating sensitive data, access to sensitive data will be restricted to authorized personnel only, using password-protected systems and, where necessary, multi-factor authentication.

A collaboration agreement that describes who can have access to what data in the project is set.

6. Ethics

Data we produce

For the data we produce, the ethical aspects are as follows:

- **All data produced during the project that does not contain personal or sensitive information will be made openly available. This includes: (i) Evaluation Data from Questionnaires: insights from researchers and supporting staff on their interest in education, internships, shadowing, mobilities, and their prior experiences with grant programs and project management. (ii) Information on Grant Programs and Calls: information regarding available grant opportunities and application processes. (iii) Information on Relevant Events and Workshops: details on upcoming workshops, educational courses, and events relevant to the research community** There are no legal, contractual, or voluntary restrictions that prevent the sharing of these datasets, as they do not contain any sensitive or personal data. If, during the project, certain data is identified as restricted due to unforeseen legal or ethical reasons, access will be limited accordingly, and these conditions will be clearly specified.
 - It does not contain personal data.

Data we collect

There are no significant ethical or legal barriers anticipated that would impede the sharing of data generated by this project. However, the project team is committed to adhering to the highest ethical standards in the collection, handling, and sharing of data.

The project involves human participants for the collection and validation of data in WP1 and WP2. To safeguard the privacy of these participants, all collected data will be automatically anonymized during the data collection process. This means that no person-specific information will be traceable, and individual responses will be aggregated and generalized to protect the identity of participants. Anonymization ensures that the data collected do not fall under the scope of the General Data Protection Regulation (GDPR), as no personal data will be processed or stored.

Participation in all questionnaires is entirely voluntary. Participants are fully informed about the purpose of the data collection, the voluntary nature of their participation, and the measures taken to anonymize their responses. The project will not include any sensitive or personal data in public repositories. All data shared will be appropriately anonymized and reviewed to ensure compliance with ethical standards.

While no ethical or legal issues are currently anticipated, the project team will remain vigilant and responsive to any changes in ethical standards or legal regulations that may impact data sharing in the future.

The data collection is not subject to ethical legislation.

7. Other issues

We use the [Data Stewardship Wizard](https://upce.dsw.elixir-europe.org/wizard) with its *Common DSW Knowledge Model* (ID: dsw:root:2.6.8) knowledge model to make our DMP. More specifically, we use the <https://upce.dsw.elixir-europe.org/wizard> DSW instance where the project has direct URL: <https://upce.dsw.elixir-europe.org/wizard/projects/23674043-1147-4798-a725-ea530ddfe401>.

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

- **Action University Development Plan**
https://www.upce.cz/sites/default/files/global/2024-01/8017/Action%20Plan_2023_2025_87451.pdf
- **Rules of the Quality Assurance and Evaluation System of Educational**
https://www.upce.cz/sites/default/files/public/onsr3560/pravidla_sys_zajist_a_ho_d_kvality_vzdel_tv_a_souvis_cin-cs-en-mar-05_rev_eng_161626.pdf
- **Creative and Other Related Activities**
https://www.upce.cz/sites/default/files/public/anle3805/upce_-_prirucka_en_-_mentoring_175605.pdf
- **Guidelines to mentoring programmes within the University**
https://www.upce.cz/sites/default/files/statutes_of_upa_commercialization_board_sm_1-2019_eng.pdf
- **Directive to intellectual property rights**
https://www.upce.cz/sites/default/files/protection_of_intellectual_property_rights_sm_1-2014_aktual_2019_eng.pdf
- **Measures to adjust the distribution of income from the commercialisation of research results**
https://fes.upce.cz/sites/default/files/public/mika0267/scipap_43_114050.pdf
- **Guidelines for authors of texts and publications**
https://ff.upce.cz/sites/default/files/public/lavi0532/submission_guidelines_final_174332.pdf
- **Instructions of circulation of internal documentation; Filing and Shredding Regulation**
- https://knihovna.upce.cz/sites/default/files/public/veko3567/smernice_c_5-2020_-_knihovni_rad_final.translated_148666.pdf

To facilitate smooth collaboration and data sharing, efforts will be made to harmonize data management practices across partner institutions, ensuring consistency in the application of these internal regulations. All internal procedures will be aligned with the Horizon Europe framework, ensuring that data management practices meet the standards for open access, data protection, and ethical considerations required by the funding body.

The project will include regular reviews of the DMP to ensure that all data management practices remain compliant with evolving institutional policies and external regulations. Any

necessary adjustments will be made to the project's data management plan to reflect changes in these guidelines.

This Data Management Plan is not a fixed document. It will be updated during the lifetime of the project based on the implementation of research activities according to individual work packages and new research data. The project will follow the guidelines on Data Management in Horizon Europa.