Twinning excellence on organic soil amendments effect on nutrient and contaminant dynamics in the subsurface

TwinSubDyn

Deliverable D5.2 Data Management Plan

Written by TwinSubDyn project consortium

N°	Short	Beneficiary	Role
1	UNSPMF	University of Novi Sad Faculty of Sciences	CO
2	UNIVIE	University of Vienna	BEN
3	FZJ	Forschungszentrum Jülich	BEN
4	MLU	Martin-Luther-Universität Halle-Wittenberg	BEN
5	CSIC	Spanish National Research Council	BEN

Place and Date: Novi Sad, Republic of Serbia 11.01.2023.

General Information

Project: TwinSubDyn

GA Number: 101059546

Call identifier: Twinning Western Balkans HORIZON-WIDERA-2021-
                            ACCESS-02

Topic: HORIZON-WIDERA-2021-ACCESS-02-01

Start date of project: 01/08/2022

Duration: 36 months

Work Package: WP5

Type: Deliverable

Number: D5.2

Title: Data Management Plan

Due Date: 31/01/2023

Prepared by: University of Novi Sad Faculty of Sciences

Responsible Person/s: Snežana Maletić

Dissemination Level: Public

Dissemination Level		
PU	Public	**X**
CO	Confidential, only for members of the consortium (including the Commission Services)	
CON	Confidential, only for members of the Consortium	

Version History				
Version number	Date	Reasons for release	Responsible	Comments
1.0	10.01.2023.	Deliverable	Snežana Maletić	Deliverable drafting
1.1	17.01.2023.	Deliverable	Gabriel Sigmund	Editing and Revision
1.2	18.01.2023.	Deliverable	Arthur Gross	Editing and Revision
1.3	23.01.2023.	Deliverable	Gordana Vlahović	Editing and Revision
1.4	25.01.2023.	Deliverable	Snežana Maletić	Final version

Table of Contents

1 SUMMARY

This deliverable describes the Data Management Plan (DMP) and includes all necessary processes and procedures which cover the full life cycle of scientific data and research results generated during the project as well as the nature of the data, how they are collected, stored and shared. The DMP guarantees that the data follow the FAIR data principle, meaning that the data are findable, accessible, interoperable and re-usable. TwinSubDyn DMP will be evolving as the project progresses and updated towards the end of the project.

2 INTRODUCTION

The Data Management Plan (DMP) describes how the data generated during the project will be collected and processed, during and after the project closure. It also explains the detailed methodology and standard procedures applied.

Basically, the DMP describes the data management life cycle for the data to be collected, processed and/or generated during the project. It also ensures adherence to the guidelines of the European Commission on FAIR Data Management in Horizon Europe as described in article 17 and analyzed in the Annotated Grant Agreement, article 17.

Among others, following information should be included:

- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the origin/provenance of the data, either generated or re-used?
- What types and formats of data will the project generate or re-use?
- What is the expected data size that will be generated or re-used?
- To whom might the data be useful ('data utility'), outside project?

3 DATA SUMMARY

Table 1 summarizes datasets, linked with the TwinSubDyn tasks, which will be collected through the project lifespan.

Table 1. TwinSubDyn project data summary

Data origin	Data formats	Purpose of the data & relation to WPs:
Reports, presentations, agendas, questionnaire, interviews, list of attendees, presentations, photos, audio visual materials, methodology used for raw data collection are standard acquisition methods by means of various experiments conducted in lab, compiled promotional materials. posters, website. Google analytics, scientific papers, TwinSubDyn social media analytics, etc.	Descriptive (.PDF, .PPTX or .DOCX), Images (.JPEG, etc.), video material (.MP4, .AVI, etc.), html links, Document (.PDF, .DOCX, .XLS, .CSV), Photos (.JPEG, .TIF etc.), Data may also come from other domain-specific databases, spreadsheets and other data formats.	Reports, presentations, agendas, questionnaire, interviews, list of attendees, presentations, photos, audio visual materials are common to tasks 1.2 -1.4 (short term scientific exchanges, on site trainings, hybrid meetings, webinars etc.); tasks 2.1 2.4 (workshops, summer school, meetings, hybrid meetings, presentations, lectures, etc.) and task 2.5 (action plan and Scientific strategy) and are included in WPs 1-2 The methodology used for raw data collection are standard acquisition methods by means of various experiments conducted in lab, photos etc. are common to tasks 1.2, and 3.1-3.6 (research tasks in the field of OSA impacts on SOM dynamics and on nutrient and contaminant sorption, transport and degradation by combining hydrogeological modelling and state of the art analytical techniques) and are included in WP3. The promotional materials. posters, website. Google analytics, scientific papers, TwinSubDyn social media analytics, etc. are common to tasks 1.2, and 3.1-3.6 (dissemination and communication activities) and are included in WP4.

Expected size of data: Storage space required for this project will be estimated between 1 and 5 TB and will be stored on a project internal volume for every partner, especially created project MS Teams platform, and UNSPMF own cloud customized for TwinSubDyn needs.

Usefulness: TwinSubDyn data (existing, new and metadata) are useful for other researchers who wish to replicate the results by means of experiment replication and validation. The obtained data through TwinSubDyn project are useful to the academic and general public working in the field of organic soil amendment application.

4   FAIR DATA

The TwinSubDyn consortium will work to ensure that its data are adhering to the 'FAIR Data Principles', which means they are Findable, Accessible, Interoperable and Reusable[1]. These aspects are herewith described:

4.1 Making internal data findable

During the lifecycle of the project, all data will be stored and systematically organized in a database at an institutional server. Data can be accessed and amended via a password protected office PC and archived at the institutional server. Individual files will be uploaded to the UNSPMF MS Teams OneDrive platform to be shared among the project team members. Secure data usage will be provided by passwords and user authentication before data accessing, downloading, and uploading. TwinSubDyn obtained a valid MS Teams license through UNSPMF, and the unique coordinator's account. TwinSubDyn section on MS Teams OneDrive platform is divided in customized subfolders (currently: 01_Templates and visibility; 02_Agreement; 03_Meetings, 04_Minutes of the meeting, 05_Deliverables; 06_Risk management; WP1_Networking for excellence, WP2_Road map for UNSPMF; WP3_Strategic research project; WP4_Dissimination and communication). Coordinator is responsible for defining the permission rights to folders for all partners in the project. Each WP leader is responsible

---

[1] https://www.go-fair.org/fair-principles/

for arrangement of the subfolder dedicated to specific WP, whereas the coordinator is responsible for the overall arrangement of folders 01 to 06. The selected data will be additionally backed up on the UNSPMF Cloud: https://cloud.pmf.uns.ac.rs/login.

Expected external findability of TwinSubDyn data

Dataset related to the WP1, 2 and 4 in Table 1 (training materials, agendas produced etc.) in TwinSubDyn will be open and findable on the TwinSubDyn website (https://twinsubdyn.pmf.uns.ac.rs/) or the TwinSubDyn knowledge platform (https://knowledge-hub.pmf.uns.ac.rs/), and will be identifiable and trackable by means of a persistent Uniform Resource Locator (URL). Only the datasets referring to the scientific strategy of UNSPMF will be restricted and available exclusively to the consortium members and findable in the Institutional database (MS Teams and UNSPMF Cloud).

Dataset related to the WP3 and task 1.2 refers to the data obtained in the research tasks which remain restricted until published in the peer reviewed scientific journals. Scientific papers will be published under Gold (>4), or Green Open Access and in agreement with EC guidelines. All the data used will be made findable and accessible in machine readable formats as supporting data files complementing the publications. The following apply for research data management:

1. Access to research results will be facilitated by publishing datasets of interest, including such data that may not be published, so as publications preprints (depending on the journal policy) in open and trusted repositories such as
   - https://www.open.uns.ac.rs/ (primarily), and/or
   - https://phaidra.univie.ac.at/, and/or
   - https://juser.fz-juelich.de/, and/or
   - https://digital.csic.es/, and/or
   - ZENODO repository (https://zenodo.org/)

   UNSPMF as TwinSubDyn project coordinator has access to an internal University of Novi Sad's open online research data repository called BE-OPEN

(https://www.open.uns.ac.rs/) where the generated and collected data will be deposited. BE-OPEN is completely harmonized with the FAIR data principles. This repository allows the researchers to deposit both publications and data, providing tools to linking them to these through persistent identifiers and data citations. Repository is set up to facilitate findability, accessibility, reusability and interoperability of data sets, which are the basic principles that the ORD pilots must comply with. BE-OPEN has enabled collecting, archiving, handling and sharing larger volumes of data in agreement with recommendations of OpenAIRE (https://guidelines.openaire.eu/en/latest/) and CERIF (https://www.eurocris.org/cerif/main-features-cerif) standards, defined by euroCris[2] organization. Advanced options implemented in BE-OPEN repository are: automatic creation of researcher profile, import of metadata from SCOPUS base using API keys, merging duplicated records, checking publishers permission by means of SHERPA/RoMEO (http://sherpa.ac.uk/romeo/) service, assignment of Creative Commons licenses, etc. The DSpace-CRIS platform enables searching entities across many criteria (keywords, authors, institutions, projects, etc.). Additionally, it is possible to observe achievements of a researcher, their network of contacts and Altmetric score (which demonstrates publication reaching in various services and popular social media).

In addition, TwinSubDyn beneficiaries can use the other online research data repositories given above, depending on data types and formats generated. Therein, not only publication related data, but more comprehensively research data and associated meta-data have their own DOI identifier. Any chosen online repository needs to facilitate identification of data and refer to standard identification mechanisms (unique identifiers such as Digital Object Identifiers – DOI; ORCID; Scopus ID; PubMed ID)

2. Researchers involved in TwinSubDyn will create personal persistent author identifiers such as ORCID, that provide a persistent digital identifier, distinguishing

---

[2] https://eurocris.org/

the researcher from every other contributor and supporting automated linkages among all professional activities.

3. Specific filters, based on the metadata elements, will allow to refine the search across datasets (e.g. by keywords).

4. Machine readable metadata will follow a standardized format, in line with community standards, and will provide information on the publication/data (author(s), publication title, date of publication, publication venue); Horizon Europe or Euratom funding; grant project name, acronym and number; licensing terms.

5. Each Task leader will be responsible for depositing the relevant data in appropriate open access online repositories.

6. Data will be made accessible within one month after publishing the data in peer reviewed scientific articles or similar, unless beneficiaries have outlined justifiable reasons for maintaining the data confidentiality.

4.2 Making data accessible

In order to maximize the impact, the project will facilitate sharing of results and deliverables within and beyond the project team. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations/posters at conferences, as well as through open access data repositories. There will be an open access policy applied to these, following the rules outlined in the Grant Agreement.

Task leaders will collect the data from each WP and the Steering committee (SC) will review and approve all the data identified as appropriate for open access. This will be an ongoing process to facilitate timely publication of the appropriate data. All data will be made available for verification and re-use, unless the task leader can justify why the data cannot be made openly accessible. The SC will assess such justifications and make the final decision, based on but not limited to: (a) IPR issues; (b) security or privacy reasons; (c) conflicts between open access rules and the national and European legislation (e.g. data protection regulations) etc.

Expected levels of accessibility of TwinSubDyn data

Dataset related to WP1, 2 and 4 in Table 1 (training materials, agendas produced etc.) in TwinSubDyn will be open and findable on the TwinSubDyn website (https://twinsubdyn.pmf.uns.ac.rs/) or the TwinSubDyn knowledge platform (https://knowledge-hub.pmf.uns.ac.rs/), and will be identifiable and locatable by means of a persistent Uniform Resource Locator (URL). Some personal data have to be restricted in accordance with the GDPR (i.e those obtained from interview or questionaries). Video materials collected during the project will be additionally uploaded on the TSD YouTube Channel but will certainly include a disclaimer on personal consent of the participants. Only, dataset referring to Scientific strategy of UNSPMF will remain confidential and restricted since it represents an internal strategic document for further development of the UNSPMF in the TwinSubDyn field.

Dataset related to the WP3 and task 1.2 referring to the obtained data in the research tasks will be restricted until published in scientific journals. Scientific papers will be published under Gold (>4), or Green Open Access and in agreement with the EC guidelines. Immediate open access through the repository listed above will be provided either to the author final manuscript (incorporating all revisions following peer-review, known as Author Accepted Manuscript or AAM) or to the final published version, known as the Version of Record (VoR). Open repository, data will be shared with other researchers also through direct correspondence, e.g. request for full publication through professional social networks such as ResearchGate and LinkedIn;

TwinSubDyn data contained in scientific publications will be made available via open-access online platforms, unless subject to protection, or if the release of all or part of the data to open-access platforms will be against the consortium interests and conditions laid down in the Consortium Agreement/Grant Agreement etc., including the data related to exploitation.

The most important results of the TwinSubDyn project (dissemination level: public) will be also accessible through the European Commission services such as Cordis EU research results (https://cordis.europa.eu/project/id/101059546).

4.3 Making data interoperable

TwinSubDyn respects the common data and metadata standards and formats as a key aspect for technological and semantic data operability. The metadata code list register (https://inspire.ec.europa.eu/metadata-codelist), will be consulted by the TwinSubDyn consortium. Standardization makes the data discoverable, and this is how the international and interdisciplinary access to, and use of the research data are promoted. To ensure the correct use and interpretation of the TwinSubDyn data by the consortium and re-users, the consortium will ensure that data is converted to common formats, allowing data exchange and reuse between researchers, institutions, organizations, countries, etc.

4.4 Making data reusable

The TwinSubDyn project will generate novel data and knowledge through experimental approaches that will be presented to the scientific community, and general public, through designed dissemination, communication and exploitation activities. Datasets uploaded in the BE-OPEN repository or other ones will be freely accessible (after an embargo period determined per dataset). Potential users are expected to adhere to BE-OPEN Terms of Use. The users of the TwinSubDyn data should acknowledge the project and/or offer co-authorship to the investigators who collected the data. Before the data are shared for re-use, appropriate Intellectual Property Licensing is implemented following the Grant Agreement Article 16.

After the end of the TwinSubDyn project, datasets available in the BE-OPEN or other open repository can be freely used by Third Parties. That means whenever possible, datasets will be licensed under an Open Access license.

For the scientific publications the latest available version of the Creative Commons Attribution International Public Licence (CC BY) or a licence with equivalent rights; for

monographs and other long-text formats will be used. The licence may exclude commercial uses and derivative works (e.g. CC BY-NC, CC BY-ND). CC-BY is the most accommodating of licenses offered and recommended for maximum dissemination and use of licensed materials because it lets others distribute, remix, tweak, and build upon licensed work, even commercially, as long as the authors are credited for the original creation.

Beneficiaries (or authors) will retain sufficient intellectual property rights to comply with the open access requirements.

Metadata of deposited publications will be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (in particular machine actionable) and will provide information at least about the following: publication (author(s), title, date of publication, publication venue, doi); Horizon Europe funding; grant name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action. Where applicable, the metadata will include persistent identifiers for any research output, or any other tools and instruments needed to validate the conclusions of the publication.

Publishing in peer-reviewed scientific journals will be used as the main channel of public disclosure of collected/generated data. All researchers will ensure that the data underlying peer-reviewed publications will be retained and made available for verification purposes after publication which is in accordance with the accepted codes and standards. Whenever suitable, researchers will aim to publish in Open Data Journals and after publications underlying data will be uploaded to appropriate public and institutional repositories. All data to be disseminated through peer-reviewed publications will be made available at the time of publication or after a prescribed embargo period at the latest (up to 6 months after publication) by the journal if green open access of the paper is implemented.

5. ALLOCATION OF RESOURCES

Consortium partners will utilize open access of peer reviewed scientific publications through free options: deposition of the accepted manuscript in author's personal web pages, institutional publication repositories and free public repositories through self-archiving (also referred to as 'green' open access). In cases when journals chosen for optimal dissemination of scientific results (timely and wide dissemination leading to significant benefit to the community) are not offering green open access or an 'embargo' period is too long, TwinSubDyn consortium will opt for the 'gold' open access.

To ensure Open Access after the lifetime of the TwinSubDyn project, the consortium will, whenever possible, use free tools that are compatible with the requirements by the OpenAIRE, such as the deposition of data and publications in online free of charge repository that allows for free and long-term deposition.

6. DATA SECURITY

For the duration of the TwinSubDyn project, datasets will be stored on the responsible partner's storage, UNSPMF MS Teams One drive and UNSPMF Cloud (access to the last 2 is restricted by user account and password). Every partner is responsible to ensure that the data are stored safely and securely and in full compliance with the European Union data protection laws.

7. ETHICAL ASPECTS

TwinSubDyn will adhere to all of the given guidelines on data management. There are no further ethical aspects foreseen within the TwinSubDyn project.

8. CONCLUSION

Data Management processes, procedures and tools have been set up for the TwinSubDyn project to establish an appropriate and efficient management as well as to ensure that the standard requirements are met. The DMP is a live document and will be updated as the project evolves, but certainly towards the end of the project.

9. ACKNOWLEDGEMENTS