



This project is part of the PRIMA programme supported by the European Union

PRIMA-SECTION 2-2022

“Modelling and Technological Tools to Prevent Surface and Ground-Water Bodies from Agricultural Non-Point Source Pollution Under Mediterranean Conditions”

NPP-SOL

Data Management Plan

Deliverable number: D6.2

Version 1.1



This project is part of the PRIMA programme supported by the European Union

Project Acronym: NPP-SOL

Project Full Title: Modelling and Technological Tools to Prevent Surface and Ground-Water Bodies from Agricultural Non-Point Source Pollution Under Mediterranean Conditions

Call: PRIMA Section 2 2022

Topic: PRIMA Section 2 2022- Multi-topic 2.1.1 RIA

Type of Action: Multi-topic Topic 2.1.1 RIA (Prevent and reduce land and water salinization and pollution due to agri-food activities).

Author(s)	A. Coppola, S. B. M. Hassan, N. Maio
Dissemination level	Sensitive
Approved by reviewers	
Date of approval	
Approved by project coordinator	A. Coppola
Date of approval	
Call	PRIMA Section 2 2022- Multi-topic Topic 2.1.1 RIA (Prevent and reduce land and water salinization and pollution due to agri-food activities).
Project website	https://mel.cgiar.org/projects/npp-sol

Copyright and legal notice:

The views expressed in this document are the sole responsibility of the authors and do not necessarily reflect the views or position of the European Commission and PRIMA Programme. Neither the authors nor the NPP-SOL Consortium are responsible for the use which might be made of the information contained in here.

Table of Contents

Executive summary	4
1. Introduction	5
1.1 Context	5
1.2 Objectives and Extent.....	5
1.3 Document Organization.....	6
2. NPP-SOL project data overview	6
3. FAIR and Open Science principles	8
3.1 Fair Data Principles.....	9
3.1.1 Findable Data.....	9
3.1.2 Accessible Data.....	9
3.1.3 Interoperable data.....	10
3.1.4 Reusable data	10
4. Handling and safeguarding personal data within NPP-SOL	10
4.1 Data Protection Officer	12
Conclusion	13
Annex I - Information Sheet & Informed consent.....	14

Executive summary

This document details the NPP-SOL Data Management Plan (DMP), as outlined in section 2.2 of the project proposal Part B. The DMP outlines the nature of data collected or generated by the project, including its purpose, responsible parties for collection and processing, and methods of utilization, storage, and transfer.

Furthermore, the DMP presents the consortium's strategies, policies, and procedures regarding data collection, storage, access, sharing, protection, retention, and disposal throughout and after the project's conclusion. It will specify whether data will be shared or made openly accessible and outline the procedures for curating and preserving data, in alignment with the European Guidelines on FAIR (Findable, Accessible, Interoperable, Reusable) Data Management. The DMP covers both research data and peer-reviewed publications, with partners committing to making all academic publications open access as per the project proposal.

However, this version of the document does not include information regarding the Intellectual Property (IP) generated by the project or an overview of relevant Intellectual Property Rights (IPR) and protection. Such details will be incorporated into subsequent versions at months 18 and 36.

1. Introduction

1.1 Context

The European Commission (EC) outlines criteria for a DMP, which includes details on managing research data both during and after project completion, specifying the data to be collected, processed, and generated, as well as the methodologies and standards to be applied. Additionally, it addresses whether data will be shared or made openly accessible, and how data will be curated and preserved post-project.

This document comprehensively covers these aspects and more. In adherence to the EC's principle of *'as open as possible, as closed as necessary'*, the NPP-SOL consortium intends to make all project-generated data openly accessible, unless there are compelling reasons to restrict access. Furthermore, NPP-SOL partners aim to make data available for future research efforts.

The NPP-SOL consortium underscores the importance of robust data management, recognizing its pivotal role in enhancing research quality. Thus, the consortium is committed to producing well-organized, documented, preserved, and accessible data, ensuring its validity is maintained throughout. This commitment is not only beneficial for the consortium itself but also for other projects seeking to utilize data from the NPP-SOL project.

The NPP-SOL data management policy is underpinned by two fundamental principles: (1) Maximizing the accessibility of generated research data in a responsible manner. (2) Ensuring that the release of data does not compromise the integrity of the research process.

1.2 Objectives and Extent

The aim of this deliverable remains consistent in showcasing the data generated by the NPP-SOL project and outlining its utilization. Additionally, it elucidates the adherence to FAIR principles (Findable, Accessible, Interoperable, Reusable) concerning data usage. Furthermore, it delineates the strategies for safeguarding and ensuring the security of data, especially personal data, produced within the NPP-SOL project.

All phases of the data life cycle are addressed, which include:

1. **Planning Research:** Involves designing research, planning data management, outlining consent procedures for data sharing, designing data collection and processing protocols, and exploring existing data sources.
2. **Data Collection:** Includes the collection of data, including both data and metadata, and acquiring third-party data if needed.
3. **Data Processing and Analysis:** Involves tasks such as data entry, digitization, transcription, translation, validation, cleaning, anonymization, creation of derivative data, documentation, storage, analysis, interpretation, and citing of data sources.
4. **Data Publication and Sharing:** Includes establishing copyright, generating user documentation, creating discovery metadata, determining access protocols, publishing, sharing data, and promoting its accessibility.

5. Data Preservation: Activities involve migrating data to optimal formats and media, storing and backing up data, creating preservation documentation, and maintaining and curating data for long-term accessibility.
6. Data Reuse: Encompasses secondary analysis, follow-up research, research reviews, scrutinizing findings, and utilizing data for educational purposes¹.

Although NPP-SOL deals with all aspects of the data life cycle as a consortium, not all partners are required to carry out every activity.

1.3 Document Organization

After this introductory section, Section 2 provides an overview of the research datasets that would be produced by the project. Section 3 delineates the project's adherence to obligations concerning open data, notably the FAIR principles. Section 4 outlines the handling and safeguarding of personal data within the project. Annex 1, for collecting consent for the use of personal data, completes this first version of the DMP.

2. NPP-SOL project data overview

The table below lists the datasets that will be compiled or generated throughout the project's activities, including the associated tasks and the responsible partner ("lead"). While several of these datasets are currently in the planning stage and do not yet exist, future updates on data management will furnish additional details.

Table 1 NPP-SOL Datasets overview

Lead	Task	Dataset name	Data utility
CIHEAM-IAM & UNIBAS	Task 4.1	Integrated modelling dataset	The main data requirements and databases structures for integrated Modelling Tools (MT)
UNICA & UB	Task 3.1	CS preliminary characterization	To identify the main NPS pollutants in the four Case Studies (CS) -Pollutant Source Analysis (PSA)
UNICA	Task 3.2	BR in Italy, Morocco and Israel	<ol style="list-style-type: none"> 1. Monitoring the chemical-microbial parameters in the bioreactors 2. Checking the hydraulic properties of the filling material
UB	Task 3.3	CW in Spain	To validate the efficiency of the constructed wetlands in Spain CS
ARO	Task 3.4	AD in Israel	To monitor the general parameters of the Anaerobic Digesters (AD) and the digestate

Table 2. Data to be collected for NPP-SOL

Dataset name	Concise list of data to be collected within the project
--------------	---

¹ <https://ukdataservice.ac.uk/learning-hub/research-data-management/>

<p>Integrated modelling dataset</p>	<p><u>Agrohydrological modelling data:</u></p> <ul style="list-style-type: none"> - pedological information; - hydraulic and hydro-dispersive properties in the different horizons of each of the available soil profile; - climatic series; - cropping pattern; - information on vegetation (LAI, crop coefficients, reference evapotranspiration, root depth, phenological phases); - information on field irrigation systems and irrigation distribution techniques; - irrigation volumes; - quality of water resources used for irrigation; and - fertilizer types, quantities and management <p><u>Bioeconomic modelling data:</u></p> <ul style="list-style-type: none"> - Meteorological data: net radiation and daily minimum and maximum temperatures - Revenues from farming activity and other activities - Self-consumption rate - Family labour - Hired labour - Seasonal labour - Working calendar (working needs per operation within the year) - Irrigation costs - Capital Investment - Cost of all the inputs for crop activity and animal activity (per input unit not the total) - Herd dynamics - CAP payments
<p>CS preliminary characterization</p>	<ul style="list-style-type: none"> - Spain CS (at Lerma gully + wherever possible also in shallow groundwater): <ul style="list-style-type: none"> -Overall water quality characterization: physico-chemical parameters; major ions, minor and trace elements, dissolved organic carbon concentration. -New isotopic characterization of Nitrogen species to confirm the sources of pollution. - Israel CS (Upper Nahalal stream watershed at the Model Farm for Sustainable Agriculture, Newe Ya'ar Research Center) <ul style="list-style-type: none"> - Overall water quality characterization: runoff, subsurface flow, irrigation by treated wastewater - nitrate and phosphorus - Livestock manure (from the local feedlot): organic carbon, nitrogen and phosphorous - Italy CS (Arborea Area in Sardinia): <ul style="list-style-type: none"> -Overall water quality characterization: physico-chemical parameters; major ions, minor and trace elements, dissolved organic carbon concentration.

	<ul style="list-style-type: none"> -New isotopic characterization of Nitrogen species to confirm the sources of pollution. - Morocco CS (Mnasra Region in the Gharb irrigated study area): <ul style="list-style-type: none"> - Overall water quality characterization: physico-chemical parameters; major ions, minor and trace elements, dissolved organic carbon concentration. - Salinity levels in the water and in the soil profile
BR in Italy, Morocco and Israel	<p>Italy and Morocco:</p> <ul style="list-style-type: none"> - Hydraulic gradient - Porosity and hydraulic conductivity of the woodchips - Inflow and outflow discharges - Inflow, intermediate and outflow pollutant concentrations - Inflow, intermediate and outflow dissolved oxygen <p>Israel:</p> <ul style="list-style-type: none"> - Hydraulic gradient - Porosity and hydraulic conductivity of the woodchips - Inflow and outflow discharges - Inflow, intermediate and outflow pollutant concentrations - Inflow, intermediate and outflow dissolved oxygen - Meteorological parameters (radiation, evapotranspiration, air temperature rainfall)
CW in Spain	<ul style="list-style-type: none"> - Inflow, intermediate* and outflow physico-chemical parameters - Inflow, intermediate* and outflow pollutant (nitrate) concentrations - Inflow, intermediate* and outflow organic carbon concentrations - Inflow, intermediate* and outflow major anions and cations - Inflow, intermediate* and outflow nitrate isotopic composition - Inflow and outflow water flow rate and hydraulic retention time <p>* Intermediate points consider the characterization of these parameters before and after the section of the CW covered by gravel.</p>
AD in Israel	<ul style="list-style-type: none"> - Inflow and outflow organic loading volumes - Inflow and outflow dissolved organic carbon - Inflow and outflow total nitrogen and phosphorous - Inflow and outflow salinity (EC, SAR) - Inflow and outflow pH and alkalinity - Outflow biogas production and characterization - Temperature monitoring

3. FAIR and Open Science principles

The NPP-SOL project adheres to the FAIR (**F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable) and Open Science principles to make NPP-SOL outputs efficient, reliable and responsive to scientific and societal challenges.

3.1 Fair Data Principles

The "FAIR (Findable, Accessible, Interoperable, and Re-usable) Data Principles" formulate principles that must be fulfilled by sustainable re-usable research data. Research data infrastructures providers should implement these principles accordingly as part of the services they offer. The principles refer to three types of entities: data (or any digital object), metadata (information about that digital object), and infrastructure.

This section outlines how the NPP-SOL project will fulfil these principles².

To make data FAIR, they must be:

- Findable, which means that data has logical and easy-to-follow rules in place to enable data to be found and that they can be easily searchable by members of the public are interested in the project.
- Accessible, which means that as many people can access and use the data as possible.
- Interoperable, which means that data can easily be exchanged and used by different partners.
- Reusable, which means that data is licensed so future researchers can use it for subsequent research.

3.1.1 Findable Data

Making data easily available is useful for future researchers and the public. They are more likely to understand the project and use its results if they can easily access the different documents.

There are several ways to make NPP-SOL data findable: the project website, scientific publications and hosting the data on repositories and archives. As a minimum, links to these different sources of data files and other relevant metadata will be listed on the project's data sharing landing page.

3.1.2 Accessible Data

NPP-SOL intends to publish the accessible (non-confidential) research results on the project website. The timing of the publication, whether immediately after submission to the PRIMA Foundation or after approval by the PRIMA Foundation, is subject to agreement.

Scientific publications play a crucial role to make research and research data findable and accessible. Metadata for such publications must include a statement acknowledging PRIMA Foundation funding, such as “*Co-funded by PRIMA (Partnership for Research and Innovation in the Mediterranean Area), Section 2 (Projects funded by participating states), Multi-topic 2022 – Topic 2.1.1 RIA (Prevent and reduce land and water salinization and pollution due to agri-food activities), NPP-SOL project, grant number, publication date, length or embargo (if applicable), and a persistent identifier*” and clarifying that the content is the responsibility of the authors, not necessarily reflecting the views of the PRIMA Foundation or of the National granting authority. This acknowledgment will be included in all publications resulting from NPP-SOL.

Open access articles are required to be deposited in a repository promptly, without undue delays. NPP-SOL has pledged to produce 8 – 10 scientific publications, all of which will be open access, and

² Mark D. Wilkinson et al. - The FAIR Guiding Principles for scientific data management and stewardship, Nature Scientific Data, 2016, 3(160018), pg.1

ensure that data is curated and preserved in line with the H2020 Guidelines on FAIR Data Management (2016), even after the project's conclusion. The NPP-SOL team will follow the procedures for prompt publications in peer-reviewed, green, and gold open-access journals. Nonetheless, beneficiaries (or authors) must maintain adequate intellectual property rights while adhering to the open access mandates.

Utilizing stable and trusted repositories is essential for ensuring the accessibility of research data, even beyond the project's duration. NPP-SOL will store data, results, reports, and publications generated by the project in publicly available data repository.

The uploading of data will occur either upon approval of the deliverables by the PRIMA Foundation or upon editorial acceptance for publication, with the deadline set at the end of the project. Each shared dataset will be assigned a persistent identifier, typically a Digital Object Identifier (DOI), along with relevant metadata and links to the project name.

NPP-SOL partners will diligently identify appropriate repositories for hosting open data sets. Potential repositories for hosting NPP-SOL datasets are outlined in the following table:

Table 2 NPP-SOL Data Repository

Repository	Description
CIHEAM-IAMM repository	https://cloud-nppsol.iamm.ciheam.org/index.php/s/coa6WaGdsjBJokG

3.1.3 Interoperable data

Data must be able to be combined and used together with other data or tools. The data format must therefore be open and interpretable by various tools, including other databases. The concept of interoperability also applies to metadata. For instance, metadata should use a standardised language shared internationally by different indexing services. In this project, it will be in English.

3.1.4 Reusable data

Data that fulfil the above conditions of findability, accessibility and interoperability are typically, therefore reusable. The NPP-SOL consortium will ensure that any product that can be made public is made public on the project website, once accepted by the PRIMA Foundation and, where appropriate, in open access repositories. Data open to the public will be created in common and appropriate file formats, respecting applicable data standards. Full re-usability of data must be accompanied by adequate documentation to ensure that the data are understandable and re-usable. A user unfamiliar with the data or the project must be able to understand what it is. Depending on the format of the data produced, the NPP-SOL consortium will discuss and document the best ways to make the data reproducible through open software repositories, source code and documentation.

4. Handling and safeguarding personal data within NPP-SOL

This section provides details on handling personal data as an integral part of the project's broader data management strategy. 'Personal data' refers to any information concerning an identified or identifiable

natural person ('data subject'). An identifiable natural person can be identified directly or indirectly, including by a name, identification number, location data, IP address, online identifier, or other factors specific to the individual's physical, physiological, genetic, mental, economic, cultural, or social identity³.

In this section we are referring to the personal data of persons who will be involved in project activities. We are not referring to project PIs and their respective teams, already officially involved in the project.

Personal information

Personal information will be solely utilized for research objectives, where it will be integrated with other data during various research endeavours aimed at achieving the goals of the NPP-SOL project. The project outcomes will be disseminated through reports to funding bodies, articles in scholarly publications, and presentations at conferences, while also being shared with the research team involved in the project.

Anonymity and confidentiality

The names and contact information provided will be used solely for research records and for sharing research outcomes. Individuals involved have the right to request the removal of their data from our records at any time. All provided information will be handled confidentially and anonymized. The NPP-SOL project will not reveal any details that could identify external participants, and explicit permission will be sought for using quotes, or they will be anonymized.

Consent form

To effectively manage human participants in research, it is essential to maintain records of participant consent, especially for identifiable data subjects. This includes preserving the informed consent form (refer to Annex 1), as well as the participant's name, contact details for coordinating their involvement, organizational affiliation, and job role for selecting suitable participants based on recruitment criteria and any specific access needs for physical events to ensure the participant's safety and comfort. Whenever feasible, participant responses are anonymized to protect privacy. This information should be securely retained by the UNIBAS and should not be disclosed. The legal basis for processing this data is the data subject's consent, as indicated in the consent forms provided in Annex I below.

Storage restriction

The principle of storage limitation mandates that personal data should not be retained in an identifiable format or for longer than required. Hence, partners only retain personal data for the duration necessary. When identifying information becomes unnecessary, it will be securely disposed of. X

³ <https://gdpr-info.eu/art-4-gdpr/>

4.1 Data Protection Officer

According to the General Data Protection Regulation (GDPR), certain organizations handling personal data are required to designate a Data Protection Officer (DPO). Public authorities and other organizations whose core activities involve processing operations that necessitate regular and systematic monitoring of data subjects on a large scale are mandated to have a DPO. However, other organizations have the option to appoint a DPO if they choose to do so.

The project coordinator is strongly advising partners to contact their DPO and inform him/her about personal data processing activities within the project. A list with the contact details of the relevant DPO will be provided in the next version of the DMP.

It has been decided for the time being to appoint one contact person per partner to whom all data management issues can be addressed.

Partner	DMP Referent contact
University of Basilicata	Antonio Coppola Shawkat B.M. Hassan
University of Cagliari	Stefania Da Pelo
Universitat de Barcelona	Manuela Barbieri Rosanna Margalef Marti
Mediterranean Agronomic Institute of Montpellier	Aybike Bayraktar
Agricultural Research Organization – Volcani Institute	Roy Posmanik
Ministry of Agriculture and Rural Development	Roey Egozi
National Institute of Agricultural Research	Abdelmjid Zouahri
Mohammed V University of Rabat	Souad El Hajjaji

Conclusion

A comprehensive plan for ensuring accessibility of NPP-SOL output data to an external audience during and after the project is currently under discussion. This discussion is not only at the level of the project partners. The aim of the project is to take this reflection and decision-making to the local level. The DMP will be relayed to the local hubs in the four case studies, allowing the primary stakeholders and beneficiaries of the initiative to contribute their perspectives on data utilization and storage.

The consortium will develop the IP strategy in cooperation with the DPO and more details will be provided in the next versions of the DMP.

Partners are also advised to discuss any new data sources with their DPO, or person responsible for data protection at their organisation.

Where partners need assistance with data protection, the Project Management Board is available to discuss issues with them.

Annex I - Informed consent

Terms of consent	Yes	No
I confirm that I have read the information sheet explaining the project and have had the opportunity to ask questions about the project.		
I confirm that my participation is voluntary and consent to data collected from me to be used for an NPP-SOL deliverable, which will be publicly available in open access repositories.		
I consent to the activity I participate in, and any data will only be used for the project's purpose.		
I consent to have the name of my organisation stated for this research.		
I consent to the project team contacting me, if required, as a follow-up to the research/engagement activity.		
I consent to project researchers' use of my anonymised responses		
I consent to the use of my name and responses in project deliverables.		
I consent that my picture is taken, and video recorded during the event (if applicable) and used for dissemination activities or others within the context of the project.		
If necessary, I consent to audio recording for future content analysis, exclusively by the project team and for the project purposes.		

Participant's full name:

Signature:

Date: DD/MM/YYYY