SUSFORAGE

Sown forage mixtures for sustainable agroecosystems in the Mediterranean area

DATA MANAGEMENT PLAN

October, 2023

Programme Call:	PRIMA Call 2020 Section 2 Topic 2.2.1. (RIA)
Project Number:	3043
Project Title:	SUSFORAGE – Sown forage mixtures for sustainable agroecosystems in the Mediterranean area.
Partners:	CTFC (Project Coordinator), CREAF, HU, UOB, KIS, INRAE
Work-Package:	WP7
Deliverable #:	D7.1
Deliverable Type:	Document
Contractual Date of Delivery:	XXXX
Actual Date of Delivery:	XXXX
Title of Document:	Data Management Plan (DMP)
Author(s):	Gemma Arnau
Content of this report:	The Data Management Plan describes the management procedures and standards for data collection and generation, security, budget, and ethical issues. It also presents the current metadata catalogue of the project.
Availability:	This report is public

Document revisions		
Author	*Revision content*	*Date*
Gemma Arnau	D7.1_v01 – First draft of the data management plan (DMP) and of the metadata catalogue	January, 27 2023
Ma. Teresa Sebastià	D7.1_v01 – Revision	October, 10 2023

TABLE OF CONTENT

INTRODUCTION

The overall objective of the SUSFORAGE project is to determine technical and socioeconomic opportunities and barriers on the transition from monocultures to diverse sown forage crops in mixed cut-grazed (polycropping-livestock) sustainable farming systems across a wide range of climates, to develop locally adapted performance models under climate change conditions in the Mediterranean.

In particular, the specific objectives of SUSFORAGE are:

- Assessing the delivering of goods and services of sown forage mixtures, including yield, nutritious feed, soil diversity and fertility, water use efficiency, pest regulation, climate change mitigation, and crop stability and resilience against climatic uncertainties.

- Designing sustainable polycropping-livestock systems to increase farming competitiveness and understanding the effects of grazing on goods and services provisioning by sown forage mixtures.

- Investigating the feeding value of forages from the species-rich meadows and how to incorporate those into the diets for livestock.

- Developing mathematical models to determine the best regionally adapted composition and proportion of sown species to guarantee resilience and stability of highly nutritious, locally produced, forage in the face of a changing climate.

- Identifying technical and socioeconomic opportunities and barriers to the establishment of polycropping-livestock systems by involving local stakeholders in case-study regions. This multi-actor approach includes gender factors, to achieve a holistic agro-livelihood system re-design.

It is expected that beneficiaries make the research data generated by the project findable, accessible, interoperable, and reusable (FAIR). The aim of the Data Management Plan (DMP) is to guarantee the deployment of best practices in data management to ensure the security and availability of data. The DMP describes the management of the data generated within the project, from its collection to treatment, storage and sharing processes, always according to the FAIR principles while considering the General Data Protection Regulation (GDPR). The DMP is designed to be a living document during the project, reviewed and updated every 6 months by CTFC, acting as DMP manager, in collaboration with WPL and other partners.

DATA SUMMARY

D.7.1 is the SUSFORAGE Data Management Plan (DMP) and is designed to set out key operational procedures to handle and manage research data, to ensure data security and quality as well as to foster data exchange and cooperation following the FAIR (findable, accessible, interoperable, and re-usable) principles. The purpose of the DMP is to support the data management life cycle for all data that will be collected, processed, or generated by the project. Indeed, the DMP describes the policy and the procedures concerning the acquisition, storage, classification, management, protection, and distribution of project data.

The document is structured after the European Commission guidelines in the following way: Section 1 is the Data Summary and describes the purpose behind data collection, process and generation, and their relation to SUSFORAGE objectives, while Section 2 explains how the FAIR principles are applied to the project, including tools, methodology and licenses adopted. Section 3 presents how the project resources are allocated for data management, while Section 4 is dedicated to data security and storage. The final sections (5-6) are dealing with the ethical and legal aspects, and further issues that may arise.

The DMP will evolve during the project lifetime. Next versions will refine the data policy aspects and will present the datasets collected and generated by the SUSFORAGE project in more detail. It is worth specifying that the SUSFORAGE Data Management Plan does not replace by any means the contractual obligations among partners, and between partners and the European Foundation.

Data management turns out to be very important for the successful achievement of the SUSFORAGE goals. At this early stage, it is important to assess the nature and extent of the datasets the project will collect, process, and generate, even if in a preliminary way. This initial version of DMP and metadata catalogue serves as fundamental step for model setup and integration for all the involved partners, to foster discussions and implementation, to then timely and successfully fulfil the project tasks.

1.1 Data Inventory Registers

The Data Inventory Register aims to collect information about the datasets collected or produced during the project. The DMP manager is the person responsible for maintaining the DIR, while the completion of the information for each dataset is the responsibility of the partner involved in the collection and use of each dataset. The information provided must be as rich as possible, and metadata fields must be filled where applicable. The metadata elements (Table 1) are included in the DIR and some of the information collected can then be used when publishing the datasets. The data generated by the project (Table 2) includes the list and the descriptive information for specific datasets and will be updated with each release of the DMP.

Table 1. SUSFORAGE Data Inventory Register metadata elements

Metadata elements	Description
Partner name	The name of the partner
Task-Subtask number	The identification number at the task or subtask level
Responsible party	The person(s) responsible for the creation and maintenance of the dataset.
Dataset name	A descriptive relative short name for the dataset.
Data Type	Example: experimental, observational, simulation Data from surveys, climate, biophysical data, etc.
Personal data	If the dataset contains information relating to an identified or identifiable natural person. Yes or No.
Sensitive data	Any data reveals a subject's information. Namely, racial, or ethnic origin, political and religious beliefs, biometric data, sexual orientation, and others. Yes or No.
Date	Date of creation of the dataset
Number of data files	The total number of files that make up the dataset
Source	Where the data originates. If data are reuse, best practice is to identify the related resource.
Format	The file formats.
Size	The total weight of the dataset (when applicable)
Dataset duration	The extant or time taken to play or execute the dataset, when applicable.
Mode of data collection	The method(s) used to collect data
Data collection instrument	The instruments used to generate and or process the data.
Support documentation	Whether documents have been created to provide context for the data (readme file)
Update frequency	The frequency with which the data will be updated
Storage location	Place where the data is stored
Preservation	Backup periodicity and for how long the dataset need to be preserved
Availability	Private, consortium or open
Rights and restrictions	Data licence CC BY-SA 4.0 is recommended for data sharing.

Table 2. SUSFORAGE Data Inventory Register metadata elements

Task	Dataset	Descriptive information

2. FAIR DATA

This section describes practices on how the project intends to make data Findable, Accessible, Interoperable and Reusable. It focuses on the approach to adopt for the project data than can be openly shared without restrictions, although the solutions described enable data protection measures.

2.1. Making data findable

SUSFORAGE will manage data through an internal server, hosted by CTFC, where all technical information (officially released documents, contractual information, templates, meeting minutes, etc.) about the project are stored in a structured way. Data are also stored and exchanged among project partners using the same server. The access to the internal repository is password-protected and restricted to designated project partners. In case of need of higher storage space, a cloud service external to the project (Dropbox Business) can be made available by CTFC.

The final project outcomes, deliverables and scientific publications will be published and maintained on the SUSFORAGE website and eventually on [Repositori UdL](#) (Universitat de Lleida), on which will be automatically become visible on Recolecta, the national aggregator of open access repositories. It will be internally discussed whether to store some data in open access on CORA.RDR as well and to which extent, leaving the final decision to the Principal Investigators (PIs) of each WP.

In addition, SUSFORAGE would adopt the [CORA.RDR repository](#) to publish project outcomes and datasets, if applicable. Using this repository, all the public data of the project will be provided with a Digital Object Identifier (DOI) and a common dataset of metadata (based on Dublin Core). Versions of each dataset would be numbered and report main and minor changes.

Not all data needs to be deposited for sharing and preservation purposes, so it is up to each beneficiary that act as main data responsible to select which data, and at what point in its life-cycle, should be deposited. Moreover, datasets can be deposited with "private" visibility, in order to respect eventual embargo periods.

2.2 Making data accessible

Data generated in SUSFORAGE could be published in the CORA.RDR repository, together with associated metadata, if applicable and agreed among the involved institutions, and if relevant for validation of scientific publications and/or deliverables. Unique identifiers, such as the Digital Object Identifier (DOI), needs to be assigned, providing a stable and consistent way to locate both the data and the metadata. The DOI for the data that will be deposited in the CORA.RDR data repository will be generated by the service.

2.3 Making data interoperable

Data sharing within and outside the project must include metadata, along with the needed documentation for others to understand and reuse the data. The SUSFORAGE project covers several disciplines and scientific areas; therefore, it is of highest priority to find a common language between partners and stakeholders, to integrate data and information from the different domains. CORA.RDR repository follows appropriate metadata standards, specifically Dublin Core. Metadata records based on Dublin Core standard ensure that project data is more easily findable and interoperable.

For the deposit of a dataset in CORA.RDR a metadata template will made available, by the DMP manager, to project partners of SUSFORAGE. This template is to be filled by the data responsible before the deposit (Table 3):

Table 3. Metadata elements in the CORA.RDR data repository

Repository field	Definition	Notes
Host Dataverse	Name of the instance where the dataset is deposited.	By default
Dataset template	Institutional template with metadata already established	If your institution has a predefined template, select it to automatically fill in some fields.
Title	Name by which the resource is known	When the title of the dataset matches the title of the related publication add the expression: "Replication data for" in front of the title.
Author	Principal investigators involved in the production of the data in order of priority. It can be a personal or corporate/institutional name.	Repeatable
Name	Full name of the creator	Personal names must follow the form: *Surname, Name*. For institutions, indicate the developed name of the institution.
Affiliation	Organization with which the author is affiliated	Use the developed name of the institution. When there is a double affiliation, the different institutions must be separated by commas. Note: Many institutions have their own regulations for institutional affiliation.
Identifier Scheme	Name identifier scheme (ORCID, ISNI, DAI, etc.)	If the *Identifier field* is used, the *Identifier Scheme* field is required.
Identifier	Unique identifiers of natural or legal persons, according to various schemes.	Format depends on the scheme. For ORCID, use the form XXX-XXX-XXXX-XXX
Contact	Person/s or institution responsible for the dataset with whom users can contact.	Repeatable
Name	Full name of the contact	Personal names must follow the form: *Surname, Name*. For institutions, indicate the developed name of the institution.

Affiliation	Contact affiliation	Use the developed name of the institution. When there is a double affiliation, the different institutions must be separated by commas. Note: Many institutions have their own regulations for institutional affiliation.
Email	Email address of the contact (this data will not be accessible to users)	This field is not exported in any schema.
Description	Abstract describing the purpose, nature and scope of the dataset.	Repeatable
Text	Abstract that explains the contents of the dataset, as well as the purpose, nature and scope of the dataset.	In case there is a related publication, the description of the dataset must not be the same as the summary of that publication. A good description is one that identifies the content of the dataset and help the user determine whether it can be used. HTML tags can be used.
Date	Date of description	YYYY-MM-DD.
Subject	Dataset knowledge area	Repeatable
Keyword	Keywords that describe important aspects of the dataset.	Repeatable
Term	Keyword that is indexed and ranked for the purpose of retrieving the dataset.	
Vocabulary	Keyword that is indexed and ranked for the purpose of retrieving the dataset.	It is recommended that the first letter of the word be capitalized. For expressions with more than one word, capitalize only the first word. You should use controlled vocabularies of your discipline or general ones.
Vocabulary URL	Link to general vocabulary	Fill in this field when using keywords from controlled vocabularies.
Related publication	Publications related to dataset data	Repeatable
Citation	Full citation of the publication	Follow the citation style of each discipline. In case the related article is not yet published, fill in the field with this information.
ID Type	Type of identifier used in the publication	DOI, handle, etc.
ID Number	Identifier of the publication	In the case of DOIs, indicate only the prefix and suffix.
URL	Link to the publication	Include the entire URL
Notes	Important additional information about the dataset that did not appear in the description.	Free text. HTML tags can be used.
Depositor	Person or Organization that deposited the dataset in the repository.	Personal names must follow the form: *Surname, Name*. For institutions, indicate the developed name of the institution.
Deposit date	Dataset deposit date in the repository	This field will be automatically filled with the current date (YYYY-MM-DD)
Kind of data	Description of the resource type.	Lists of controlled values.

2.4 Making data reusable

Research data and software are owned by the PI (beneficiary) that generates them. In case of joint ownership of results, each PI must agree (in writing) on the allocation and terms of exercise of their joint ownership, stipulating a joint ownership agreement, to ensure compliance with their obligations. The PRIMA Foundation may assume ownership of results to protect them (art. 26.4 of GA for further details). Notwithstanding the above, owners of open results arising from the SUSFORAGE project are encouraged to release their work under a Creative Commons license, preferably Creative Commons Attribution 4.0 CC-BY-4.0 license.

Concerning the dissemination and exploitation of results, each PI must 'disseminate' its results by disclosing them to the public by appropriate means (i.e., peer-reviewed journal articles) and ensure open access – free of charge, online access for any user. Further, any dissemination of results must indicate that it reflects only the author's view and that the PRIMA Foundation is not responsible for any use that may be made of the information it contains. Any update of relevant project outcome will be published on the SUSFORAGE website.

In regard to research data and results ownership, each PI may transfer ownership of its results. Nevertheless, the PRIMA Foundation has the right to object to any transfers or licensing. Quality assurance concerning accuracy and completeness of metadata will be performed with mandatory participation and collaboration by the other partners, since they are responsible for data collection, process, and generation within their tasks – in agreement with the project ethical obligations.

Any update concerning data collection and/or generation within the SUSFORAGE project as well as any editing of the metadata catalogue should be promptly communicated to the Project Coordinator and to the principal investigators by short email notice.

3. ALLOCATION OF RESOURCES

Costs are included in the tasks related to data collection and generation and cannot be listed separately. Costs of the project internal repository are covered by internal XXX resources. As already mentioned, Dropbox Business accounts will also be activated if needed, in order to have more space where to store and share raw datasets and simulation outputs. Eventual costs related to these accounts will be changing according to the numbers of accounts and disk space needed for each of them and they will be covered by project budget.

4. DATA SECURITY

The internal project repository on the CTFC server is protected by firewall and institutional security policies. In detail:

- The internal repository is relying on CTFC storage facilities and accessibility is reserved, protected by username and password known only by the selected users (PIs), for both upload and download functionalities.

- CORA.RDR repository is hosted by Consorci de Serveis Universitaris de Catalunya (CSUC) and it is subject to its rules for data security as reported at CORA.RDR Technical Features. CORA.RDR uses an open-source software called Dataverse, developed by Harvard's Institute for Quantitative Social Science. It allows you to store datasets and display descriptive metadata and downloading the corresponding files.

- All datasets maintained on the CTFC server will be periodically subject to incremental backup in order to avoid data loss.

Finally, in regard to security, SUSFORAGE does not involve any activity raising security issues and does not handle EU classified information, neither as background nor as result.

5. ETHICS

SUSFORAGE will involve stakeholders that will participate in a series of meetings both in person and online. The purpose is to bring their local and regional expertise into the project and to share knowledge and develop capacity in the approaches and tools used in the project. Confidentiality will be assured, in that:

- No personal information will be collected from participant stakeholders other than contact information and a brief description of their work, and this information will not be shared beyond the limits of the project without the explicit written consent of the stakeholders.

- No quotes made by the stakeholders in the online discussion of the workshops will be published in print or on the internet without their explicit written consent.

- No further ethical issues are anticipated, but should they arise, the project consortium will ensure that EU legislation, international guidelines and the ethical and legal requirements of the countries involved in the project are adhered to.

Ethical aspects related to data management are following the obligation to comply with ethical and research integrity principles for which the PIs must follow the ethical principles (including the highest standards of research integrity) and the applicable international, EU and national law. The PIs must ensure that the activities, also the data-related ones, under the action have an exclusive focus on civil applications. In addition, the PIs must respect the fundamental principle of research integrity:

- Reliability in ensuring the quality of research data.

- Honesty in developing, undertaking, reviewing, reporting, and communicating research data in a transparent, fair, and unbiased way.

- Respect for colleagues, research participants, society, ecosystems, cultural heritage, and the environment.

- Accountability for the research data from idea to publication, for data management and organisation, and well as for their wider impacts meaning that the PIs must ensure that persons carrying out research tasks and processing data follow the good research practices and refrain from the research integrity violations.

Finally, informed consent will come along with data sharing and long-term preservation, in case of questionnaires dealing with personal data implemented within the project.

6. OTHER ISSUES

At the time of the DMP deployment, the SUSFORAGE project does not make use of any other national/funder/sectorial/departmental procedures for data management. This will be reported in case of occurrence on a later stage of the project (further versions of the DMP).

This DMP has been created with eiNa DMP, a specific tool provided by the Consortium of University Services.