D7.2 Data management plan
30/06/25

Lead beneficiary: Pensoft Publishers

Author/s: Nikolay Mehandzhiyski, Bilyana Avtova

LAFERIA D7.2 DMP

Prepared under contract from the European Commission
Grant agreement No. 101181492
EU Horizon Europe Research and Innovation Action

Project acronym: LAFERIA

Project full title: LAndscape FEatures Reintroduction in Intensive Agricultural land

Project duration: 1 January 2025 - 31 December 2028 (48 months)

Project coordinator: Francisco Moreira, BIOPOLIS

Call: Horizon-CL6-2024-BIODIV-01-7

Deliverable title: Data Management Plan

Deliverable Nº: D7.2

WP responsible: 7

Nature of the deliverable: Report

Dissemination level: Public

Lead partner: Pensoft Publishers

Recommended citation: APA

LAFERIA D7.2 DMP

Due date of deliverable: 30 June 2025

Actual submission date: 30 June 2025

Deliverable status:

Version	Status	Date	Author(s)
1.0	Draft	24.06.25	Bilyana Avtova, Nikolay Mehandzhiyski
1.1	Draft revision	26.06.25	Francisco Moreira, Angela Lomba, Marta Neto, Guy Pe'er
2.0	Draft	27.06.25	Bilyana Avtova, Nikolay Mehandzhiyski
2.1	Draft revision	30.06.25	Francisco Moreira, Angela Lomba, Marta Neto, Guy Pe'er
3.0	Final version	30.06.25	Bilyana Avtova, Nikolay Mehandzhiyski

Table of Contents

Key takeaway messages

- Deliverable D7.2 Data Management Plan (DMP) provides an overview of LAFERIA's data, and other expected research outputs, anticipating the way they will be generated and re-used, as well as how these project resources will be managed, abiding to FAIR principles.
- The DMP explores the ways in which generating and reusing data will contribute to LAFERIA's mission, describing the formats, licenses, standards, and data repositories whose use is planned.
- D7.2 provides a set of data management recommendations for all LAFERIA partners to follow upon the collection, generation, or re-use of their research data or other research outputs expected in the project with the objective of upholding FAIR principles, while also considering the use of artificial intelligence, and other ethical aspects of data management.

1 Summary

To ensure the findability, accessibility, interoperability, and reusability of its research data, and other project outputs, LAFERIA has developed a tailored Data Management Plan (DMP), that constitutes Deliverable 7.2. The DMP builds upon the principles of open science, open data, and open sources, pre-established in the project's grant agreement, as well as insights provided by the consortium through a survey circulated ahead of the creation of the DMP. The document begins with a concise overview (Chapter 4) of the expected collected (Table 1), and the re-used (Table 2) data. The document opens with an Introduction (Chapter 3) followed by a concise overview (Chapter 4) outlining the expected data to be collected (Table 1) and the data to be reused (Table 2). Additionally, the DMP outlines the FAIR principles, provides guidance on their implementation by LAFERIA partners in managing the data, and offers specific recommendations (Chapter 5).

Finally, the plan dives into the administrative side of data management, such as resource allocation (Chapter 6), data security (Chapter 7), intellectual property rights (Chapter 8), and ethics consideration (Chapter 9). The DMP is a living document that will be updated as needed to reflect more detailed information as the project evolves and key developments arise. Currently, two DMP updates are planned, one due in month 24, and other due in month 48.

2 List of abbreviations

EU	European Union
DMP	Data Management Plan
DoA	Description of Action
TG	Target Group
AI	Artificial Intelligence
API	Application Programming Interfaces
DOI	Digital Object Identifier
DPO	Data Protection Officer
EBV	Essential Biodiversity Variable
EML	Ecological Metadata Language
FAIR	Findable, Accessible, Interoperable, Reusable
GBIF	Global Biodiversity Information Facility
GDPR	General Data Protection Regulation
ORE	Open Research Europe
WP	Work Package
LF	Landscape Features
JRC	Joint Research Centre
CS	Case Study/Case Studies
GIS	Geographical Information System

3 Introduction

As a research-focused project which intends to produce policy, and business recommendations for the reintroduction of landscape features in agricultural areas, LAFERIA strives to embody the principles of open science, and findable, accessible, interoperable, and reusable (FAIR) data. These pillars foster knowledge sharing, enhance reproducibility, and enable new discoveries through data-driven collaboration. In alignment with these principles, the Horizon Europe project LAFERIA supports transparency and collaboration in scientific research by making its data openly accessible to both the scientific community and the broader public.

In order to have a tangible overview of these principles, including action points, LAFERIA's work package (WP) 7 is responsible for the creation of a Data Management Plan (DMP; Deliverable 7.2) in month 6 of the project. The DMP outlines LAFERIA's approach to data collection, processing, and management throughout the project, as well as its long-term preservation strategy, ensuring that any project-generated data is easily discovered, accessed, and reused by both humans and machines. The plan describes the type of data expected to be generated by the project, accompanied by the preferred licences, methodologies, standards, and metadata specifications.

The DMP was developed in accordance with the initial open science, data, and source principles outlined in the grant agreement, as well as through a

data management survey circulated amongst consortium members to accurately pinpoint their data management plans, and needs. Project members contributed to the development of the DMP by completing a comprehensive questionnaire, in which they identified the datasets and other research outputs they plan to generate or reuse, and described their institutional data management practices.

4 Data summary

LAFERIA will collect, generate and reuse data within its lifespan, in support of the following objectives:

- To assess the implementation, spatial distribution, and connectivity of landscape features (LF) across case study (CS) regions and agricultural systems, providing a clear understanding of their current status and context.
- To quantify LF coverage at the EU-wide level, identifying key patterns and regional differences to inform planning and policy development.
- To evaluate the needs and perspectives of farmers and stakeholders regarding LF reintroduction, based on qualitative data from interviews and surveys.
- To analyse the costs and benefits of LF for biodiversity, ecosystem services, climate adaptation, agricultural yields, incomes, and broader societal impacts.
- To develop tailored recommendations and implementation plans for both farmers and policymakers, grounded in regional needs and cost-benefit evidence.

LAFERIA's work will rely on a mixture of reused and self-generated data, such as qualitative data coming from interviews and surveys. Table 1 presents an overview of datasets that the LAFERIA consortium is anticipating to generate. Complementary, Table 2 outlines external data and information that the project partners are anticipating of reusing - this means external and already available data which will be used to support and enhance LAFERIA's research and activities.

The majority of the generated data is either qualitative data or spatially-explicit data (Geographical Information System - GIS - data. This reflects some of the project's core missions: i) gaining a better understanding of the occurrence and distribution patterns of LF across regions; ii) achieving a better understanding of farmers' and stakeholders' attitudes towards the reintroduction of LF; and iii) applying knowledge to inform policy recommendations.

Given that much of the data generated within the scope of LAFERIA will be qualitative and may include personal or sensitive information, particularly data collected through interviews and surveys, certain datasets will not be made fully open. In accordance with GDPR and ethical standards, datasets containing sensitive information may remain closed or be partially redacted to protect privacy, unless explicit permission is granted by the third parties involved.

Table 1 identifies the intended users of data generated by LAFERIA. In some cases, these users fall within LAFERIA's core Target Groups (TG), as defined in the project's Description of Action (DoA).

- Farmers and other land-users (TG1)
- Private and public landowners (TG2)
- Policy makers (TG3)
- Research and Academia (TG4)
- Extension services (TG5)
- Private corporations and retail (TG6)
- Administration (including agricultural chambers) (TG7)
- Civil society (landowners unions, NGOs, Community based organisations) (TG8)
- General public (TG9)
- Stakeholders in case studies and relevant initiatives (TG10)

In this case, the data users are identified as external. However, LAFERIA-generated data may also be used internally by project members collaborating across WPs. As shown in Table 2, while some of the data expected to be reused originates from external sources, other datasets will be reused internally across WPs. This means that data generated within a specific WP or task may be used by another, which may explain why certain information is not yet available for those datasets.

Table 1: Summary of the data LAFERIA consortium anticipates generating. (TBD: To be determined; N/A: Not applicable)

No	Name of the dataset	Generator(s)	WPs/Tasks	Generated via	Size	Format	Type of data	Sensitive data	Personal data	Delivery	Metadata	Users	Access	Re-use
1	Maps of the seven (7) case study regions with landscape elements	WUR	WP 1.2	Modelling and data processing	TBD	pdf, shp	GIS-data	Yes	No	M18	Standard use for GIS data	TG1, TG3, TG4, TG8	PDF will be provided, other data depends on restrictions	N/A
2	5-7 maps of landscape connectivity	WUR	T1.3	Modelling and data processing	TBD	pdf, shp	GIS-data	No	No	M28	Standard use for GIS data	TG1, TG2 TG3, TG4, TG8	Public	LARCH
3	EU-wide map	WUR	T1.2	Modelling and data	TBD	pdf, shp	GIS-data	No	No	M18	Standard use for	TG3, TG4	Public	GIS pro

	with % LF			processing							GIS data			grams
4	Interviews in case study areas	WUR	T4.2/T4.4	Interviews/survey	TBD	docx	Qualitative data	Yes	Yes	M36	N/A	LAFERIA consortium	Closed	N/A
5	Database with contacts of stakeholders in the region	WUR	T4.2/Link to WP5	Survey/contacts	TBD	xls	Qualitative data	Yes	Yes	M10-36	N/A	LAFERIA consortium	Closed	N/A
6	Inventory of initiatives	EV-ILVO	T3.2	Survey	1000 entries	xlsx	Qualitative data	Yes	Yes	M6	N/A	TG4	Closed	N/A
7	One-phone-call interviews	Project partners	T3.2	Interview	100 interviews	mp4 xlsx, or docx	Qualitative data	Yes	Yes	M12	N/A	TG4	Closed	N/A
8	EU Agricultural	BIOPOLIS	T1.1	Geostatistical analysis	TBD	shp	GIS data	No	No	M12	INSPIRE	TG3, TG4, TG7,	Public	Metadata

	systems map										TG10		
9	Landscape features coverage	BIOPOLIS	T1.2	Geostatistical analysis	TBD	shp	GIS data	No	No	M18	INSPIRE	TG3, TG10	Public	Metadata
10	Evidence synthesis	University of Helsinki	T2.1 / T2.2	Literature review	1GB	xlsx	Quantitative data	No	No	M18	Standard	TG4	Public	N/A
11	Case study data	University of Helsinki	T4.2	Interviews	500 MB	xlsx, docx	Qualitative data	No	Yes	M32	Standard	TG4	Public, whilst personal data is omitted	N/A
12	EU policy analysis	IEEP	T3.1	Extracts from plans or policies	TBD	xls	Qualitative data	No	No	M10	Date of screening of policy documents	IEEP	Internal use only (for the moment)	N/A
13	Scoping review	Halmstad University	T3.1	Systematic	TBD	pdf	Quantitative	No	No	?	?	TG4	Open	N/A

				literature review			data							
14	Initiatives	Halmstad University	T3.2	Documents, websites and dialogue	TBD	xls	Qualitative data	No	No	M6	N/A	TG1, TG2, TG3, TG4, TG5	Open	N/A
15	Case study	Halmstad University	T3.3 T4.4	Interviews, focus groups, workshop	TBD	Transcriptions	Qualitative data	No	No	M36	N/A	TG1, TG2, TG3, TG4, TG5	Closed	N/A
16	Interviews	Halmstad University	T3.3 T4.4	Interviews	TBD	Transcriptions	Qualitative data	No	No	M36	N/A	TG1, TG2, TG3, TG4, TG5	Closed	N/A
17	Costs and benefits	UROS	T2.3	Interviews, data catalogues	30 observations	xls	Cost-Benefit data	Yes	Yes	M30	N/A	WP2, WP4, T6.3	Internal use as it is sensitive data (including researchers in	Task 2.2 / Task 2.4

													Algeria)	
18	Motivation decision-making process	UROS	T4.3	Qualitative interviews	20-30 observations	MaxQDA	Qualitative data	Yes	Yes	M46	N/A	WP4, T6.3	Internal use as it is sensitive data	Task 2.2, 2.4, Task 4.2, 4.3
20	Interviews/focus group recordings and transcripts	Syke	T4.2, T4.3 / T5.2	Workshops and interviews	1Mb	docx, pdf, mp4	Qualitative data	Yes, to be used only if aggregated	Yes	M7 onwards	All metadata will include: title and description, tags and categories, the person and date of creation and the last modification	TG4	Closed (anonymised/internal use)	N/A

21	Land-use and vegetation of Finnish case study	Syke	T1.2	GIS data	TBD	shp, qgz	GIS	No	No	M7	All metadata will include: title and description, tags and categories, the person and date of creation and the last modification	TG4	Open	N/A
22	LAFERIA Stakeholder network	EEB	T5.1	Initial survey	TBD	xls	Qualitative data	No	Yes	N/A	N/A	All TG welcome	Internal use	Possibly all WP
23	Stakeholder engagement from the	UFZ	WP3, 4 / T5.2	Interview	1Mb	docx, pdf, mp4	Qualitative data	No	Yes	TBD	TBD	TG4	Anonymised	N/A

	Leipzig case study													
24	Maps of LFs	UFZ	T1.1	Data processing	10-20 Mb	pdf, .csv	GIS data	No	No	M10	LF types and character istics	All	Depen ds on permis sion	GIS
25	Model results	UFZ	T1.3, T2.4, T6.2	Modelling	100 Mb	Various (ascii, shapefi les)	Digital data, GIS data	No	No	M24. M48	TBD	TG4	Open access	[Per sef one .jl](#) mo del
26	Policy analysis results	UFZ	T4.1, T.61, T6.3, T6.4	Literature review, policy review	1Mb	docx, xlsx, pdf	Qualit ative data	No	No	M14, M40, M45	TBD	All	Open access	N/A
27	Case study informat ion	UFZ	T4.1	Literature review, policy review	1Mb	docx, xlsx, pdf	Qualit ative data	No	No	M14	TBD	All	Open access	N/A
28	Inventory of LF initiative s	UFZ	T3.3	Literature review	1Mb	docx, xlsx, pdf	Qualita tive data	No	No	M26	TBD	All	Open access	N/A

Table 2: Summary of the data LAFERIA partners expect to reuse (TBD: To be determined; N/A: Not applicable)

No	Name of the dataset	Contact	Task/WP	Size	Format	Sensitive data	Personal data	Metadata	Access	Origin	Licence
1	LUCAS	WUR	T5.1	TBD	shp	TBD	No	EU standard	Depends on agreement with owner	JRC	TBD
2	Data on LF from CS	WUR	T1.2	TBD	TBD	TBD	TBD	TBD	TBD	TBD	TBD
3	LPIS	STEP	WP1	TBD	GIS	No	No	Yes	Open	Ministry of Agriculture	TBD
4	Agricultural systems	BIOPOLIS	T1.1	10Mb	tiff	No	No	EEA standards	Closed	JRC	N/A
5	Energy input	BIOPOLIS	T1.1	7.2 Mb	tiff	No	No	EEA standards	Closed	JRC	N/A
6	Corine land cover 2018	BIOPOLIS	T1.1	125 Mb	geotiff	No	No	INSPIRE	Open	Copernicus Land Monitoring Service (EC)	CC BY
7	Evidence library JRC	UH	T2.1, T2.2	TBD	xlsx	No	No	TBD	Open	JRC	Open
8	EUROSTAT public statistics	IEEP	T3.1	TBD	Online	No	No	No	Public	EUROSTAT	Open
9	Costs for implementation of agroforestry	UROS	T2.3	TBD	Online	No	No	NO	Public	KTBL	Open

10	Costs for landscape maintenance	UROS	T2.3	650 kb	xlsm	No	No	TBD	Public	KTBL	Open
11	Kostendatei Landschaftspflege 1998	UROS	T2.3	Booklet	Paper booklet	No	No	TBD	Public	Agency for Environment Bavaria	No
12	Agricultural statistics and farm level data	Syke	Research and background for CS	20 Mb	Xls, qgz	Yes	Yes	Descriptive metadata	Restricted	Natural research Institute of Finland, Food Authority of Finland	CC BY
13	Maps for the city of Leipzig	UFZ	Various	10-20 Mb	csv, shp, geotiff	No	No	TBD	Public	Open sources / City of Leipzig	TBD
14	IACS data	UFZ	Various	TBD	csv	No	No	TBD	Restricted (depending on ownership)	Ministries of agriculture	N/A
15	Input data for modelling in Persefone.jl	UFZ	T1.3, T2.4	100 Mb	csv, shp, geotiff	No	No	TBD	Public	German Meteorological Service (DWD), Mundialis, Ministries of Agriculture,	Datenlizenz Deutschland and other open licenses

									German Cartographic Office (BKG)	

5 FAIR data

Published in 2016, the 'FAIR Guiding Principles for scientific data management and stewardship', provides guidelines to improve **F**indability, **A**ccessibility, **I**nteroperability, and **R**eusability (Wilkinson et al., 2016). The principles highlight the importance of machine-actionability. In other words, the ability of computer systems to automatically find, access, integrate, and reuse data.

This chapter provides recommendations on applying the <u>FAIR Data Principles</u> to LAFERIA's data.

5.1 F - Making data findable

Maximising the value of data begins with making it easily discoverable by both humans and machines. This involves assigning globally unique identifiers, supplying detailed metadata, and indexing both the data and its metadata in searchable repositories.

- **Recommendation #1: Identifiers**

  Use globally persistent and unique identifiers. This includes identifiers such as a digital object identifier (DOI) - persistent digital identifier ensuring a stable and reliable link to the resource, or an international standard serial number (ISSN), used to identify serial publications and ongoing integrating resources, including websites (under specific conditions). More information, as well as a list of such identifies, and services providing them, is available <u>here.</u>

- **Recommendation #2: Metadata**

  Ensure metadata is machine-readable and follows standardised formats. Include detailed descriptions of the data's context, quality, and condition, along with keywords that allow computers to automatically categorise and prioritise the data.

  → **Recommendations #2.1:** Several different suitable standards have been identified by LAFERIA consortium that can be used based on the partner's preference, and specific data needs.
    - <u>INSPIRE</u>: Enables the harmonised sharing and interoperability of spatial environmental data across Europe to support policy-making and public access.
    - <u>Ecological Metadata Language (EML)</u>: Incorporates modules covering the spatial, temporal, taxonomic, and thematic scope of the data, along with the research methods and protocols used.

- ▪ <u>Darwin Core</u>: Enables the sharing of information on biological diversity, primarily centred on taxa, their occurrences, and associated data.
- ▪ <u>DataCite Metadata</u>: Offers a standardised framework for data integration across various disciplines.

→ **Recommendation #2.2:** The metadata should clearly reference the dataset's globally unique and persistent identifier.

→ **Recommendation #2.3:** If no specific metadata standard is applied, at the very least, include the following fields:
   - ❖ Creator(s)
   - ❖ Creation date
   - ❖ Title
   - ❖ Identifier
   - ❖ Licence of use
   - ❖ Embargo period
   - ❖ Publishers
   - ❖ Publication date
   - ❖ Dataset description
   - ❖ Dataset language
   - ❖ Input data/parameters/fillers/protocols used to generate the dataset
   - ❖ Ontology-aligned keywords and meta-args
   - ❖ Version
   - ❖ Size
   - ❖ Format
   - ❖ Project acronyms (LAFERIA) and number (101181492)

- ● **Recommendation #3: Indexing**
To ensure your research (meta)data is findable online, select data repositories that register, index, and make your (meta)data discoverable.

5.2 A - Making data accessible

Once users have located the data, they need to understand how to access it. This involves reducing access barriers by depositing data in trusted repositories, ensuring (meta)data can be retrieved using a standardised communication protocol and its identifier, and maintaining access to metadata even if the data itself is no longer available.

- **Recommendation #4: Repositories, and access protocols**
  Research data should be readily accessible online through standard, free, and open protocols, rather than depending on specialised or proprietary tools or communication methods.

  → **Recommendation #4.1:** Upload datasets to trusted repositories and infrastructures, such as Zenodo or GBIF (Global Biodiversity Information Facility), which assign a DOI and enable straightforward download via a simple HTTPS link, without requiring specialised tools for access. This ensures that anyone with the link can retrieve the data using a standard web browser, provided the data is openly available.

  → **Recommendation #4.2:** If there are access conditions for the data, ensure they are clearly specified, enabling automated systems to manage them or alert users to any requirements.

  → **Recommendations #4.3:** Regarding a general-purpose trusted repository, most partners have indicated an interest in using Zenodo, as they have experience with it. Thus, when not using domain specific repositories, LAFERIA partners are encouraged to use Zenodo, as the project has confirmed that it meets all FAIR data requirements, with particular emphasis on:

- **Recommendation #5: Open data rules**
  In accordance with LAFERIA's Grant Agreement, LAFERIA will apply Open Science Practices to all methodologies, research results, deliverables, policies, and scientific publications, as well as data not subject to the General Data Protection Regulation (GDPR). LAFERIA partners can ensure that their data is open access by:

  → **Recommendation #5.1:** Depositing the data in a trusted, open-access repository under the most recent version of the Creative Commons Attribution International Public License (CC BY), the Creative Commons Public Domain Dedication (CC0), or an equivalent license. This can be done using a general-purpose repository like Zenodo, a subject-specific platform such as GBIF, or a reliable institutional repository (such as Tuhat in University of Helsinki).

  → **Recommendation #5.2:** Data supporting scientific publications should be deposited in a repository by the time the paper is published and should be made openly accessible as soon as possible, and no later than 12 months after publication. All other research data should be shared and made open at the earliest possible opportunity.

➔ **Recommendation #5.3:** Metadata for deposited datasets must be openly available under CC0 or an equivalent license, provided that legitimate interests or constraints are respected. While datasets themselves may become unavailable over time due to maintenance costs, the metadata should remain accessible. This ensures that essential information continues to be discoverable and valuable for research, including replication efforts.

LAFERIA will comply with the principle of making data 'as open as possible, as closed as necessary'. In cases where necessary, access to specific datasets may be restricted. Table 1 outlines any datasets classified as restricted or closed. When a dataset is classified as restricted or closed it is typically because within the table it is also indicated as containing sensitive information. In some cases where datasets contain sensitive/personal information but still need to be open, such information will be omitted or restricted.
Project partners will evaluate the finalised datasets to determine whether they can be shared in an aggregated and/or anonymised form to protect sensitive or personal information.

5.3 I - Making data interoperable

Once users locate and access LAFERIA's data, the datasets should ideally be formatted and structured to support seamless integration with analysis, storage, and processing tools or workflows. LAFERIA will ensure its datasets are compatible with others by using widely accepted knowledge representation languages, standard vocabularies, and by including links to relevant metadata.

- **Recommendation #6:** Employ a formal, accessible, widely used, and broadly applicable language for representing knowledge, such as the Resource Description Framework (RDF), Web Ontology Language (OWL), DARPA Agent Markup Language (DAML), or JSON for Linking Data (JSON-LD).

- **Recommendation #7:** Utilise documented and publicly accessible data standards and vocabularies such as Darwin Core and the European University Information Systems (EUNIS) to ensure clarity and consistency for all dataset users.

- **Recommendation #8:** Provide clear and accurate references to other (meta)data by indicating when a dataset builds upon another, requires additional datasets for completeness, or is supplemented by related information found elsewhere**.**

5.4 R - Making data reusable

LAFERIA data will be made openly available in the public domain under licenses that permit reuse, enabling the widest possible access and use by third parties even after the conclusion of the project. To help users determine the data's relevance to their needs, each dataset will be accompanied by standardised, comprehensive metadata. This metadata will include essential details such as the data's origin, scope, limitations, methods of generation, technical specifications, related species, and more. To avoid confusion that could hinder reuse, licensing terms will be clearly specified in both human- and machine-readable formats. Where applicable, additional documentation such as README files, software or tool guides, and annotated R Markdown scripts will be provided to support data analysis, validation, and reuse.

- **Recommendation #9:** Ensure your data has a clear data usage licence.

- **Recommendation #10:** Describe data using standardised metadata to facilitate discovery, including also README files or documentation essential for data analysis, validation, and reuse. This should cover details such as data provenance, scope, limitations, generation process, and more, including items like the following: data origin, provider, date of data retrieval, licence of use, a URL or DOI for the original data, description of the input data, including its format, structure and content, and steps taken to clean and re-process the data in order to ensure both data quality and integrity.

- **Recommendation #11:** When depositing in a data repository, the author should also include clear information on additional materials, such as research outputs, software tools, instruments or methodologies, which could be necessary if a third-party should want to reproduce, re-use or validate the original data.

6 Allocation of resources

FAIR data management involves certain costs that can be grouped into two main categories. The first includes article processing charges (APCs) for publishing in open-access journals. The second relates to potential fees for depositing data in trusted open-access repositories. While some repositories, such as Zenodo and GBIF, allow free data deposition, others, like Dryad, may require users to pay a publishing fee. With these costs in mind, adequate

resources have been allocated to LAFERIA's members to support the FAIR and effective management of LAFERIA's data and other research outputs. Each project partner is expected to use its funding responsibly, with a strong emphasis on prioritising open access practices.

7 Data security

As previously mentioned, upon the completion of LAFERIA's open access datasets, they will be deposited in trusted repositories, such as Zenodo, GBIF, European Open Science Cloud, Dryad or GitHub, which guarantee the datasets' preservation over the long term. In the interim, project partners will store incomplete datasets on their own servers, all of which have regular and/or automated backup systems in place. Access to these servers is strictly controlled and protected using measures such as two-factor authentication, VPN connections, and/or firewall security.

- **Recommendation #12:** To ensure the security of LAFERIA's data, project partners are encouraged to perform regular backup procedures and confirm the official Data Protection Officer (DPO) of their institution and ensure ongoing communication with them.

8 Intellectual property rights (IPR) management

Intellectual Property Rights (IPR) are legal protections granted for intellectual creations such as knowledge, research, and creative expressions. They play a vital role in the dissemination and use of research outcomes by giving creators control over their work, ensuring legal safeguards, and enabling recognition, commercial opportunities, and exclusive usage. IPR includes forms such as patents, copyrights, and trademarks.

IPR management in LAFERIA will be conducted in accordance with Article 16 of the Grant Agreement (GA) and Chapters 8, 9 of the Consortium Agreement (CA).

8.1. Background and access rights

'Background' refers to any data, knowledge or information, regardless of its form (physical or non-physical). This includes any associated rights, such as IPR that were in the possession of the project partners prior to their joining the Agreement, or any IPR that are necessary for carrying out the action or exploiting results. If any partner wants to use background that is subject to

rights of a third party, they must ensure that they have permission to do so, and that this permission allows them to meet their obligations under the Grant Agreement.

In Attachment 1 of the CA, partners have listed and agreed on which background will be used in the scope of LAFERIA. Partners have also informed each other should any background have legal restrictions or limitations due to reasons such as licensing or confidentiality.

All partners are required to provide access to their relevant background when it is necessary for the completion of project activities. Each party is responsible for carrying out its tasks in line with the Consortium Plan and shall be solely accountable for ensuring that its actions within LAFERIA do not knowingly violate any third-party intellectual property rights.

8.2 Ownership

'Result' refers to any tangible or intangible project outcome, including data, know-how or information in any form (written, digital, physical, etc.), whether it can be protected or not, as well as the rights attached to it, whatever they may be, including IPR.

As stated in Chapter 8 of the CA, all results are owned by the partner that has generated them. Joint ownership of results is governed by the Grant Agreement (Article 16.4 and Annex 5). Unless otherwise agreed, each joint owner may use the jointly owned results for non-commercial research and teaching without needing prior consent or paying royalties. For other uses, including granting non-exclusive licenses to third parties (without sub-licensing rights), the owner must give at least 45 days' notice to the others and offer fair and reasonable compensation. Additionally, all joint owners must agree in advance on protection measures and how to share the associated costs.

The granting authority has the right to use non-sensitive information, materials, and documents from the beneficiaries (such as summaries, deliverables, and media) for policy, communication, and publicity purposes, both during and after the action. All beneficiaries grant the granting authority a royalty-free, non-exclusive and irrevocable licence to use their materials, documents and information, including the following rights:

- The granting authority may use said materials, documents and information for **its own purposes**, including making them available to its staff or any EU service, institution, body or agency copying or reproducing them in unlimited quantities, and communicating them through press information services.

- The granting authority may use said materials, documents and information with the **purpose of distribution to the public**, including publication as hard copies and/or in electronic or digital format, publication on the internet, as a downloadable or non-downloadable file, broadcasting, public display, communication through press information services or inclusion in accessible databases or indexes.
- The granting authority has the **right to edit** or redraft said materials and documents, including shortening, summarising, inserting other elements, extracting and dividing parts.
- The granting authority has the **right to translation**.
- The granting authority has the **right to storage** in both papers, electronic or other form.
- The granting authority has the **right to archive**.
- The granting authority has the **right to authorise third parties** to act on its behalf.
- The granting authority has the **right to process**, analyse and produce derivative works.

8.3 Protection

Beneficiaries who have received funding under the grant must ensure that their results are adequately protected, if protection is possible and warranted. This protection should be maintained for an appropriate period of time and cover the relevant geographical areas. The decision to protect the results should take into account various factors, such as the potential for commercial exploitation, the legitimate interests of other beneficiaries, and any other relevant considerations that may justify or require protection.

During LAFERIA's initial stages (before month 6), partners were asked to fill in a survey providing information on the data they anticipate to be generated or re-used, as well as the preliminary plan for the licenses under which said data will be made available (Annex 1). A majority of partners have indicated that they will be using a CC BY license, which allows for the use, sharing and distribution of content as long as proper credit is given to the original creator.

9 Ethics

In all activities funded by the European Union, ethics is a key part of research at every stage, from start to finish. Adherence to ethical standards is regarded as vital. Ethical research conduct refers to the application of fundamental ethical principles and legislation to scientific research in all

possible domains. All LAFERIA conducted activities must comply with ethical principles as well as national, EU and international legislation, the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights.

All data within LAFERIA will be handled in adherence to ethical principles and security concerns. Only data that is necessary and directly related to the project's objectives will be collected and processed. When personal information is gathered during research activities, participants will be clearly informed that it will be used solely for the purposes of the project. All personal data will be handled in compliance with applicable data protection regulations. Consent will be obtained from participants involved in interviews, surveys, and workshops for the processing of their personal data, and this data will not be used for any other purpose without their additional consent. As LAFERIA deals with a great portion of interviews and surveys collecting personal, often sensitive data, all consortium members will ensure to be compliant with GDPR rules. More extended knowledge on the topic can be found in the

10 References

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). *The FAIR guiding principles for scientific data management and stewardship*. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

11 Annex: DMP questionnaire

LAFERIA Data Management Plan Survey

This questionnaire aims to collect information about the various types of data that LAFERIA members will generate or store. Based on your feedback, a Data Management Plan (DMP) will be created to: 1) document the ownership, licensing and use of the project data; 2) describe the metadata; 3) store safely and enable subsequent use of the research data. The DMP will also define the datasets to be published for open use and the chosen trusted repository. It is a living document with two planned updates in M24 and M48.

There are 13 questions in this survey (approx. ~20min) and it is vital that each partner provides as detailed a response as possible, as this will guide the project's data management practices.

☐ **To continue please first accept our survey privacy policy.**
Show policy

Next

LAFERIA D7.2 DMP

LAFERIA Data Management

We realise you may not yet have exact answers to some of these questions. Please feel free to give us your preliminary estimate, which you can modify/update on your demand when needed or during the mandatory DMP updates before the end of the project (M24 & M48).

In the current DMP, we will clarify that the information it contains is based on an initial data mapping and is subject to change in order to reflect the project's development and arising data needs.

*1. First and last name

[   ]

*2. Organisation/institution
ⓘ Check all that apply

- ☐ BIOPOLIS
- ☐ PENSOFT
- ☐ WR
- ☐ HU
- ☐ BEE/EEB
- ☐ EV ILVO
- ☐ STEP
- ☐ IEEP
- ☐ UH
- ☐ UFZ
- ☐ EDIA S.A.
- ☐ UROS
- ☐ Syke

LAFERIA D7.2 DMP

Questions 3 and 4 concern the data you will **generate**.

3. Please provide the following provisional information for your **generated data**:

1. **Name of the dataset**

2. **Name of the generator**: name of the person who will generate this data.

3. **Relevant task**

4. **Generated via**: for example, field work, modelling, data processing, remote sensing, literature review, policy review, interview, surveys.

5. **Size**: a rough estimate only if you know.

6. **Format**: for example, .docx; .xlsx; .pdf; .mp4; .xml; .csv.

7. **Type of data**: for example, qualitative data; semi-quantitative data; quantitative data; analogue data; digital data; GIS data.

8. **Sensitive data**: Yes/no. If yes, please specify, for example, racial, political, ethical, health, and more here.

9. **Personal data**: Yes/no. If yes, please specify, for example, name, surname, address, email, IP address, location data.

10. **Delivery**: a rough estimate of a timeline. If there is an embargo period, specify why and how long it will apply.

11. **Metadata**: what metadata will be accompanying your datasets.

12. **Users**: to whom they might be useful. Including but not limited to: **Farmers and other land-users (TG1), Private and public landowners (TG2), Policy makers (TG3), Research and Academia (TG4), Extension services (TG5), Private corporations and retail (TG6), Administration (including agricultural chambers) (TG7) , Civil society (landowners unions, NGOs, Community based organisations) (TG8), General public (TG9), Stakeholders in case studies and relevant initiatives (TG10)**

13. **Access**: will they be open access? If not, please indicate the reasons, for example, ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, contract.

14. **Re-use**: potential documents or tools needed to re-use or validate the data.

❓ This question concerns the data you will **generate**.
❗ Please fill in at least one answer

	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
Name of the dataset					
Name of the generator					
Relevant task					
Generated via					
Size					
Format					
Type of data					
Sensitive data					
Personal data					
Delivery					
Metadata					
Users					
Access					
Re-use					

LAFERIA D7.2 DMP

Questions 6-12 concern your **data management practices**.

*6. Please provide a brief summary of your **institutional data management practices**, specifying:

1. **Data location:** where and how data will be stored, for example, institutional server or web hosting.

2. **Server location:** EU or non-EU; compliant or not-compliant with applicable data protection rules (for example, GDPR).

3. **Backup procedures:** type of backup procedures and their frequency.

4. **Protection:** how data security is ensured, for example, password or two-factor authentication.

5. **Responsible:** name the person from your team who will bear primary responsibility for data management and serve as a contact person if questions arise.

❷ This question concerns your **institutional data management**.

Data location	
Server location	
Backup procedures	
Protection	
Responsible	

*7. Do you follow a specific naming convention? (e.g., Projectname_month_year_datasetname_version)

If yes, please specify.

❷ This question concerns your **data management**.
❶ Choose one of the following answers

○ Yes

○ No

Please enter your comment here:

31

LAFERIA D7.2 DMP

Question 5 concerns the data you will obtain from elsewhere and **reuse.**

5. Please provide the following information for the **existing data you will reuse:**

1. **Name of the dataset**

2. **Relevant task**

3. **Size**

4. **Format:** for example, .docx; .xlsx; .pdf; .mp4; .xml; .csv.

5. **Sensitive data:** Yes/no. If yes, please specify, for example, racial, political, ethical, health, and more here.

6. **Personal data:** Yes/no. If yes, please specify, for example, name, surname, address, email, IP address, location data.

7. **Metadata:** the metadata accompanying your datasets.

8. **Access:** open/restricted/closed access.

9. **Origin:** what is the origin of the data?

10. **Ownership:** who owns the data you will reuse?

11. **Licence:** under what licence can you use the data? (Similar to a software license, a data license defines how others may use data that you have created or own and have made available—such as through a data repository.) You can find out more about data licences here

12. **Re-use:** potential documents or tools needed to re-use or validate the data.

❓ This question concerns the data you will obtain from elsewhere and **reuse.**
❗ **Please fill in at least one answer**

	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
Name of the dataset					
Relevant task					
Size					
Format					
Sensitive data					
Personal data					
Metadata					
Access					
Origin					
Ownership					
Licence					
Re-use					

Questions 6-12 concern your **data management practices**.

*6. Please provide a brief summary of your **institutional data management practices**, specifying:

1. **Data location:** where and how data will be stored, for example, institutional server or web hosting.

2. **Server location:** EU or non-EU; compliant or not-compliant with applicable data protection rules (for example, GDPR).

3. **Backup procedures:** type of backup procedures and their frequency.

4. **Protection:** how data security is ensured, for example, password or two-factor authentication.

5. **Responsible:** name the person from your team who will bear primary responsibility for data management and serve as a contact person if questions arise.

❓ This question concerns your **institutional data management.**

Data location	
Server location	
Backup procedures	
Protection	
Responsible	

*7. Do you follow a specific naming convention? (e.g., Projectname_month_year_datasetname_version)

If yes, please specify.

❓ This question concerns your **data management.**
ℹ Choose one of the following answers

○ Yes

○ No

Please enter your comment here:

*8. Do you use any standard metadata vocabulary, standards or methodologies when creating your datasets? You can find the EU's guidelines for ensuring quality in open data and metadata here.

If yes, please specify.

❷ This question concerns your **data management.**
❶ Choose one of the following answers

○ Yes

○ No

Please enter your comment here:

*9. Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

If yes, please specify.

❷ This question concerns your **data management.**
❶ Choose one of the following answers

○ Yes

○ No

Please enter your comment here:

*10. How will you licence your data? You can find information on Creative Commons licences here, and general information on data licences here.

If other, please provide a justification.

❷ This question concerns your **data management.**
❶ Comment only when you choose an answer.

☐ Creative Commons Attribution International Public License (CC BY) (or equivalent)

☐ Creative Commons Public Domain Dedication (CC 0) (or equivalent)

☐ Other

34

LAFERIA D7.2 DMP

*11. Do you have a preference for a trusted repository where to store your research data?

If yes, please specify.

❷ This question concerns your **data management.**
❶ Choose one of the following answers

○ Yes

○ No

Please enter your comment here:

---

*12. Can you identify potential obstacles (e.g., technical, social, policies) that would prevent delivering FAIR data during LAFERIA's lifetime and beyond? Information on FAIR data here.

If yes, please specify.

❷ This question concerns your **data management.**
❶ Choose one of the following answers

○ Yes

○ No

Please enter your comment here:

Any additional comments?