

Deliverable 7.6 – Data Management Plan



**Co-funded by
the European Union**

LIFE23-CCM-ES-LIFE
MERLIN 101158094



Data Management Plan–M6

D7.6: Data Management Plan–M6

Summary

This report includes the initial version of the LIFE MERLIN Data Management Plan (DMP), which follows the European Commission's guidelines and the Horizon 2020 FAIR Data Management Plan Template. This plan describes the entire lifecycle of research data, from collection to preservation, including: Data collection, processing, and generation during the project; Methodologies and standards applied to ensure data quality and integrity; Data sharing and open access strategies; Curation and preservation approaches, including post-project maintenance. The DMP is a dynamic document that will evolve throughout the project, adapting to the actual research data generated and ensuring that our data management practices remain effective and efficient.

Deliverable number	Work package
D7.6	WP7
Lead beneficiary	Deliverable author(s)
CETAQUA	Marcel Vilaplana (CETAQUA)
Quality assurance	
Alicia González	CETAQUA Galicia
Planned delivery date	Actual delivery date
28/02/2025	07/03/2025 (v1.0) 26/05/2026 (Updated version v2.0)
Dissemination level	• PU = Public

LIFE MERLIN is a European project co-funded by the LIFE Programme (LIFE23-CCM- ES-LIFE MERLIN 101158094), the European Union's funding instrument for the environment and climate action. The general objective of LIFE is to contribute to the implementation, updating, and development of EU policy and legislation on environment and climate through the co-financing of projects with European added value. The opinions and views expressed are solely those of the author(s) and do not necessarily reflect those of the European Commission or CINEA. Neither the European Commission nor the granting authority can be held responsible for them.

Table of contents

List of Figures	3
List of Tables	4
List of Acronyms and Abbreviations	5
Executive summary	6
1 Objectives	7
2 Data Summary	8
3 Fair Data	13
3.1 Making data findable, including provisions for metadata	13
3.2 Making data openly accessible	14
3.3 Making data interoperable	14
3.4 Increase data re-use (through clarifying licenses)	15
4 Allocation of resources	15
5 Data security	16
6 Ethical aspects	16
7 Other issues	16
References	17

List of Figures

Figure 1. Data Management procedure to be adapted in LIFE MERLIN

8

List of Tables

Table 1. Data summary in LIFE MERLIN	10
---	-----------

List of Acronyms and Abbreviations

EC: European Commission

DMP: Data Management Plan

FAIR: Findable, Accessible, Interoperable, and Re-usable

GDPR: General Data Protection Regulation

IPR: Intellectual Property Rights

WP: Work Package

Executive summary

This report includes the initial version of the LIFE MERLIN Data Management Plan (DMP), which follows the European Commission's guidelines (European Commission (a)) and the Horizon 2020 FAIR Data Management Plan (DMP) Template. This plan describes the entire lifecycle of research data, from collection to preservation, including:

- Data collection, processing, and generation during the project
- Methodologies and standards applied to ensure data quality and integrity
- Data sharing and open access strategies
- Curation and preservation approaches, including post-project maintenance.

The DMP is a dynamic document that will evolve throughout the project, adapting to the actual research data generated and ensuring that our data management practices remain effective and efficient.

1. Objectives

The Data Management Plan (DMP) is a crucial element that outlines the data generated during project execution and its accessibility. The DMP provides a comprehensive analysis of the data management policy employed by the LIFE MERLIN project, encompassing all datasets generated or collected.

This document outlines the following essential aspects of data management:

- Data Handling: How research data will be managed during and after the project
- Data Collection and Generation: Types of data to be collected, processed, and generated
- Methodologies and Standards: Approaches and standards applied to ensure data quality and integrity
- Data Sharing and Accessibility: Strategies for data sharing and open access
- Data Curation and Preservation: Methods for curating and preserving data, including post-project maintenance

This initial version of the DMP for the LIFE MERLIN project will undergo updates at M24, M36, and M48 to reflect the actual research data generated. This document is structured according to the European Commission guidelines and the Horizon 2020 FAIR Data Management Plan template.

2. Data Summary

The Data Management procedure in LIFE MERLIN will follow the scheme included as Figure 1 (Hansen et al., 2018).

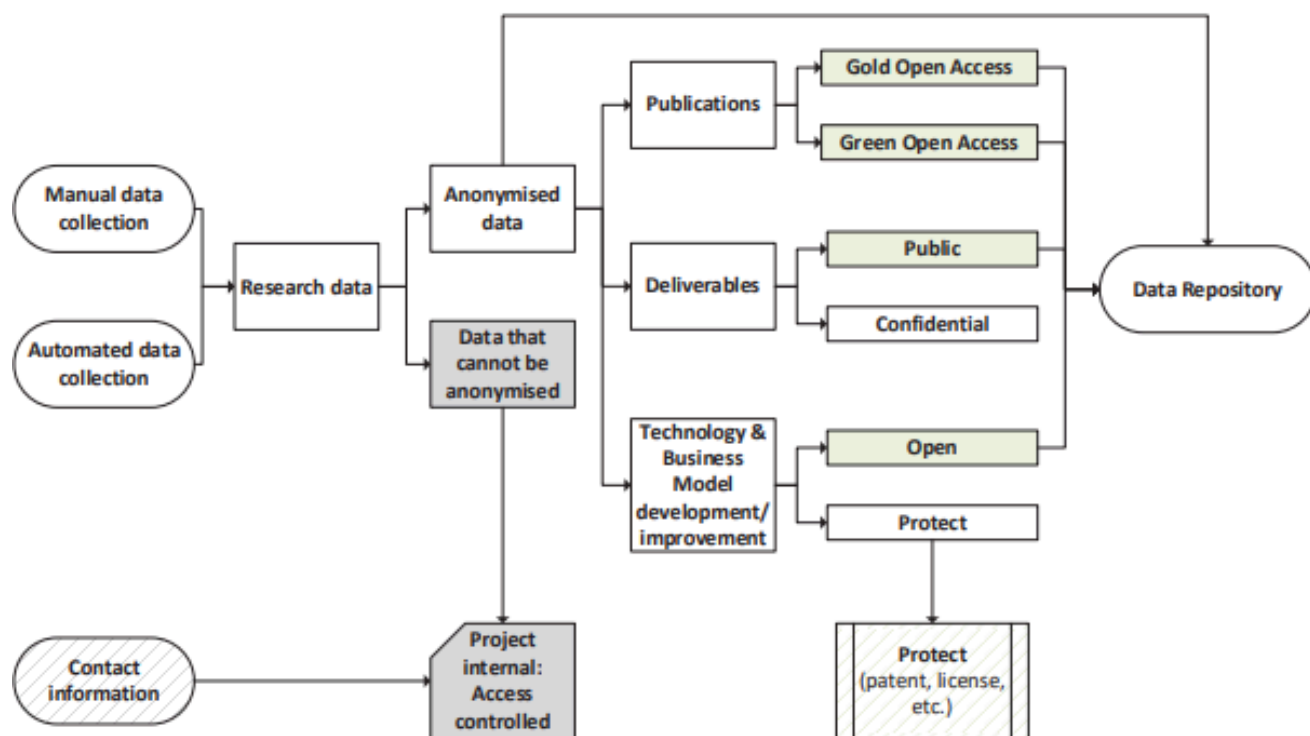


Figure 1. Data Management procedure to be adapted in LIFE MERLIN

The most important aspects of the data management are presented below:

- Research data will be gathered from various sources, including:
 - Pilot studies (e.g., operational data from demo plants, pumps, sensors)
 - Lab studies (e.g. operational data, data, analysis results)
 - Personal contact information (e.g., from the website or during workshops)
- Data Access and Protection. To ensure responsible data management, the project will
 - Implement controlled access to personal data, adhering to general rules on personal data protection.
 - Utilize non-personal data and anonymized personal data for project work, publications, and deliverables.

- Protect data with potential commercial exploitation through collaborative risk management among project partners and the coordinator.
- Publication and Dissemination: LIFE MERLIN will adopt an open access model for scientific and technical publications. Most project deliverables will have a public dissemination level and will be available on the project website. For data linked to Intellectual Property Rights (IPR) and copyright considerations, specific protection measures will be established.
- Data Repository and Access: A data repository will be established to provide access to open research data. The LIFE MERLIN website may be used to host all the project information and generated data considered non-sensitive or non-confidential.
- Confidentiality and Data Sharing: Most data generated in the LIFE MERLIN project will be sensitive or confidential. The main information will be about the design and operation of the demo plants, and will be shared only with authorized researchers and end-users.

A comprehensive overview of the data to be collected, processed, and/or generated during the LIFE MERLIN project is provided in Table 1.

Table 1. Data summary in LIFE MERLIN

Related WP	Purpose	Types and Formats	Re-use of existing data	Origin (data collected, processed or generated)	Size	Confidentiality	Data utility (Beneficiary)
WP1	Conceptual design of the demo plants	Report (PDF)	Yes	Data collected and generated	Less than 100 Mb	Public	CET/AMA/EMU/CREA
WP1	Basic Design of Murcia demo plant	Report (PDF)	Yes	Data collected and processed	Less than 1 Gb	Sensitive	EMU/CET/CREA
WP1	Basic Design of Alicante demo plant	Report (PDF)	Yes	Data collected and processed	Less than 1 Gb	Sensitive	AMA/CET/CREA
WP1	Tender Process	Contact Information	No	Data collected and processed	Less than 100 Mb	Confidential	EMU/AMA/CET
WP1	Murcia demo plant description	Report (PDF)	No	Data collected	Less than 1 Gb	Sensitive	EMU/CET/CREA
WP1	Alicante demo plant description	Report (PDF)	No	Data collected	Less than 1 Gb	Sensitive	AMA/CET/CREA
WP2	Co-digestion optimization at lab-scale	Data	No	Data generated	More than 1 Gb	Sensitive	All Partners

WP2	Development, implementation and commissioning of the digital-tool	Data / Report (PDF)	Yes	Data collected and processed	More than 1 Gb	Sensitive	CREA/AMA/EMU/CET
WP3	Operation and optimization of demo plants	Data / Report (PDF)	No	Data generated	More than 1 Gb	Sensitive	All Partners
WP4	Sustainability plan	Report (PDF)	Yes	Data collected and processed	Less than 1 Gb	Public	ACES/CET/AMA/EMU
WP4	Replication case studies	Data	Yes	Data collected and processed	Less than 1 Gb	Sensitive	ACES/CET/CREA
WP4	Business plan	Report (PDF)	Yes	Data collected and processed	Less than 1 Gb	Sensitive	ACES/CET/CREA
WP5	Impact indicators	Data / Report (PDF)	Yes	Data generated and processed	Less than 100 Mb	Public	All partners
WP5	LCA assessment	Report (PDF)	Yes	Data collected and generated	Less than 100 Mb	Public	All partners
WP5	LCC Assessment	Report (PDF)	Yes	Data collected and generated	Less than 100 Mb	Public	All partners
WP6	Communication	Report (PDF)	No	Data collected	Less than 10 Mb	Public	All partners



WP6	Dissemination activities and workshops for engaging stakeholders	Contact information	No	Data collected	Less than 10 Mb	Confidential	All partners
WP7	Project management and coordination	Personal information / Report (PDF)	No	Data collected	Less than 10 Mb	Confidential	All partners



3. Fair Data

The LIFE MERLIN project is committed to managing data in accordance with the principles of FAIR data management (Findable, Accessible, Interoperable, and Re-usable data). The project aims to maximize access and re-use of research data generated by the project, while protecting sensitive information.

To balance data sharing with privacy and intellectual property considerations, the project will make non-confidential datasets publicly available to promote re-use and collaboration. Regarding the confidential datasets, the access to them will be restricted and only authorized to the members of the consortium, and specific researchers and end-users only.

The project is exploring two options for sharing non-confidential datasets:

Option 1: Create a dedicated section for non-confidential datasets in the LIFE MERLIN project website. The metadata will adhere to project website standards, and naming conventions and versioning will be based on Harvard University standards.

Option 2: Create a LIFE MERLIN profile on the CORDIS website for non-confidential datasets. The metadata, naming conventions, and versioning will follow CORDIS website rules.

Regarding the confidential data, Google Drive and CETAQUA's server will be used to share the confidential datasets between project partners. FAIR principles will be used to ensure optimal match between project partners.

3.1. Making data findable, including provisions for metadata

The making of data findable is described for the different options previously discussed:

- Dedicated section for non-confidential datasets in the LIFE MERLIN Project website
 - Harvard University naming convention standards will be used for all the data generated in the project, including a clear version number
 - Search keywords will be available to ensure the online findability and

promote the re-use of the dataset

- LIFE MERLIN profile on the CORDIS website for non-confidential datasets
 - Naming and versioning
 - Search keywords will be provided following CORDIS rules to ensure the online findability and promote its re-use.
 - Data and metadata are registered or indexed in a searchable resource (online CORDIS' search engine).
- Google Drive and CETAQUA's server to share the confidential datasets between project partners
 - The Harvard University naming convention standards are applied with all the data generated in the project, including a clear version number
 - No search keywords will be provided because the datasets will not be public
 - Data and metadata will be registered for the Google Drive website/application

3.2. Making data openly accessible

The majority of datasets generated in the project will not be publicly accessible due to confidentiality concerns. This includes the optimization of the co-digestion process at lab scale, the design and operational data from the demo plants, as well as datasets related to the replicability of the technology.

Certain Work Packages contain non-confidential data that can be shared and will be openly accessible through the project website. This includes project management data, such as Data and Risk Management Plans, and green procurement guidelines, as well as project communication and dissemination data. Additionally, impact monitoring data, including impact indicators, LCA and LCC assessments, will also be made available.

3.3. Making data interoperable

Datasets produced in the project are generated through the collection and processing of data from lab tests, in-situ sensors from demo plants, as well as other data collected directly from the pilot. Due to the confidential nature of these datasets, access will be

restricted to internal researchers and end-users directly related to the project.

3.4. Increase data re-use (through clarifying licenses)

This section will be completed once the pilot data has been analyzed and data management options have been determined for each dataset. Due to the confidential nature of the project's data, only authorized researchers and end-users will be able to re-use data in this project.

To ensure responsible data management, the project will adhere to the FAIR data management guidelines, considering the following key aspects:

- Data Licensing: How will the data be licensed to permit wide re-use?
- Data Availability: When will the data be made available for re-use? If an embargo is sought, why and how long will it apply?
- Data Re-usability: Are the data produced and/or used in the project usable by third parties after the project's end? If re-use is restricted, why?
- Data Longevity: How long is it intended that the data remains re-usable?
- Data Quality Assurance: Are data quality assurance processes described?

4. Allocation of resources

LIFE MERLIN will utilize standard tools and a free research data repository to store and manage the data from the project. Internally, data will be stored on partners' and CETAQUA's existing servers, as well as online via Google Drive existing accounts. This approach eliminates the need for additional costs or external resources. Only non-confidential data will incur costs associated with open access portals/websites.

The costs of data management activities will be limited to protect management costs and will be covered by allocated resources in the project budget. More details on data management costs and resources will be provided in this section once a specific data management option has been selected.

5. Data security

Cetaqua's digital security policies dictate a protocol for managing and storing confidential data, including data shared and/or generated by all project partners.

LIFE MERLIN data is stored and shared across multiple sites, each requiring an appropriate level of security:

- Partner hardware (Computers and server)
- A shared LIFE MERLIN Google Drive account, created under Veolia's account, provides a secure platform for data sharing. Access to this account is strictly managed by the project coordinator and is only granted to project partners and advisory board members who have signed a non-disclosure agreement.

6. Ethical aspects

All project partners are obliged to protect personal data following European and national General Data Protection Regulation (GDPR).

7. Other Issues

No other issues have been detected



References

European Commission (a). Data Management Guidelines – H2020 Online Manual.
https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

C. W. Hanssen, M. Natvig, J. Gorman, S. Jiang, R. Søråsen, and A. Schoevaars. (2018). "Data Management Plan".

