



# SEMPRE-BIO

## D7.6 Data Management Plan V3

**SEcuring doMestic PProduction of  
cost-Effective BIOmethane**

**CETAQUA**  
WATER TECHNOLOGY CENTRE



Funded by  
the European Union

## PROJECT INFORMATION

<b>GRANT AGREEMENT NUMBER</b>	101084297
<b>PROJECT TITLE</b>	SEcuring doMestic PROduction of cost-Effective BIOmethane
<b>PROJECT ACRONYM</b>	Sempre-Bio
<b>FUNDING SCHEME</b>	HORIZON-IA
<b>START DATE OF THE PROJECT</b>	1 November 2022
<b>DURATION</b>	42 months
<b>CALL IDENTIFIER</b>	HORIZON-CL5-2021-D3-03-16
<b>PROJECT WEBSITE</b>	<a href="https://sempre-bio.com/">https://sempre-bio.com/</a>

## DELIVERABLE INFORMATION

<b>DELIVERABLE N°</b>	D7.6
<b>DELIVERABLE TITLE</b>	Data Management Plan v3
<b>WP NO.</b>	7
<b>WP LEADER</b>	CETAQUA
<b>CONTRIBUTING PARTNERS</b>	CET, AB, CRYO, DBFZ, DTU, INV, PROPULS, SINTEF, TERRA, TMB, UGE, UVIC, BIOGAS-E, INNOLAB, NAT, MASS
<b>AUTHORS</b>	Alejandra Córdova
<b>REVIEWERS</b>	Jaqueline Daniel-Gromke (DBFZ)
<b>CONTRACTUAL DEADLINE</b>	31 Oct 2025
<b>DELIVERY DATE TO EC</b>	31 Oct 2025
<b>DISSEMINATION LEVEL</b>	Public

## DOCUMENT LOG

VERSION I	DATE	AUTHOR	DESCRIPTION OF CHANGE
V0.1	04-04-2023	Alejandra Córdova	Initial version
V0.2	19-04-2023	INV (Laia Mencia)	Review feedback
V1.1	21-04-2023	Alejandra Córdova	First version
V1.2	25-04-2023	INV (Laia Mencia, Estafanía González)	Review feedback
V1.3	26-04-2023	Alejandra Córdova	Final version ready for submission
VERSION 2	DATE	AUTHOR	DESCRIPTION OF CHANGE
V2.0	08-10-2024	Alejandra Córdova	Initial version ready for review
V2.1	24-10-2024	DBFZ (Torsten Thalheim, Jaqueline Daniel-Gromke)	Review feedback
V2.2	28-10-2024	Alejandra Córdova	Complete version ready for review
V2.2	29-10-2024	DBFZ (Jaqueline Daniel-Gromke)	Review feedback
V2.2	28-10-2024	Alejandra Córdova	Final version ready for submission
V2.2	15-11-2024	Imke Lubbeke	Review feedback
V2.3	13-12-2024	Alejandra Córdova	Complete version ready for review
V2.3	17-12-2024	DBFZ (Jaqueline Daniel-Gromke, Thalheim, Torsten )	Review feedback
V2.4	16-01-2025	Alejandra Córdova	Complete version ready for review
V2.4	07-02-2025	DBFZ (Jaqueline Daniel-Gromke)	Review feedback
V2.4	12-02-2025	Alejandra Córdova	Final version ready for submission
V3.0	27-10-2025	Alejandra Córdova	Third version ready for review
V3.1	30-10-2025	DBFZ (Jaqueline Daniel-Gromke)	Review feedback
V3.1	31-10-2025	Alejandra Córdova	Complete version ready for submission

## DISCLAIMER



**Funded by the  
European Union**

“Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.”

## CONTENT

1. Executive Summary	6
2. Data Summary	7
3. FAIR Data	9
3.1. Making data findable	9
3.1.1. Data discoverability	11
3.1.2. Data identification mechanisms	11
3.1.3. Naming conventions	11
3.1.4. Document versioning	12
3.1.5. Metadata standards	13
3.2. Making data accessible	14
3.2.1. Data availability	14
3.2.2. Data access	14
3.3. Making data interoperable	15
3.4. Increase data re-use	15
4. Allocation of resources	17
5. Data security	18
6. General Data Protection Regulation - GDPR	19
7. Ethical Aspects	21
8. Other issues	22
9. History of Changes	23

## Figure index

Figure 1. Front-page of the Google Drive folder.	10
--	----

## Table index

Table 1. Summary of datasets generated within SEMPRES-BIO	7
Table 2. Document template extract showing the history panel.	12
Table 3. Document template example showing the information panel on the front page.	13
Table 4. File formats for different disciplines that are either recommended or acceptable	15
Table 3. History of changes.	18

## Acronym Glossary

<b>CERIF:</b> Common European Research Information Format	<b>DMP:</b> Data Management Plan
<b>EBIEs:</b> European Biomethane Innovation Ecosystems	<b>EEA:</b> European Economic Area
<b>FAIR:</b> Findable, Accessible, Interoperable and Reusable	<b>GDPR:</b> General Data Protection Regulation

## Consortium partners

	Participant organisation name	Acronym
1	CETAQUA	CET
2	AIGÜES DE BARCELONA	AB
3	CRYO INOX	CRYO
4	DEUTSCHES BIOMASSEFORSCHUNGSZENTRUM GEMEINNÜTZIGE GMBH	DBFZ
5	DANMARKS TEKNISKE UNIVERSITET	DTU
6	INVENIAM GROUP	INV
7	PROPULS	PROPULS
8	SINTEF AS	SINTEF
9	TERRAWATT	TERRA
10	TRANSPORTS METROPOLITANS DE BARCELONA	TMB
11	UNIVERSITEIT GENT	UGE
12	UNIVERSITAT DE VIC	UVIC
13	BIOGAS-E	BIOGAS-E
14	INNOLAB	INNOLAB
15	NATURGY	NAT
16	NV De Zwanebloem	MASS
17	VEOLIA WATER TECHNOLOGIES TECHNO CENTER NETHERLANDS BV	BIOTHANE

## 1. Executive Summary

SEMPRE-BIO (SEcuring doMestic PROduction of cost-Effective BIOmethane) is a €9.9M project financed under the Horizon Europe Cluster 5 programme running from November 2022 to April 2026. SEMPRE-BIO aims to demonstrate novel and cost-effective biomethane production solutions and pathways, deemed essential to achieve the European Green Deal and climate and energy targets for 2030 and the net zero greenhouse gas emissions by 2050, and to increase the market up-take of biomethane-related technologies.

With sites in Baix Llobregat (ES), Bourges (FR), and Adinkerke (BE), SEMPRE-BIO will establish three European Biomethane Innovation Ecosystems (EBIEs), which will be indicative of the various baseline settings for biomethane production throughout Europe. The challenge is to lower investment and operating costs, optimize feedstock supply and use, identify alternative feedstock, and reduce their costs, improve plant efficiency and operations, account for carbon savings, and increase and monetize co-benefits, such as from the commercialization of the digestate or the valorisation of residual gas streams.

The SEMPRE-BIO Data Management Plan (DMP) follows the DMP template that was designed to be applied to any Horizon Europe project that produces, collects or processes research data.

The aim of the DMP is to plan the life cycle of data within the SEMPRE-BIO project. It offers a long-term perspective by outlining how data will be generated, collected, documented, shared and preserved within the project.

The DMP is a living document to be updated as the implementation of the project progresses in the context of the periodic evaluation/assessment of the project and when significant changes occur.

## 2. Data Summary

The SEMPRES-BIO Data Management Plan (DMP) provides the final strategy for managing the data generated and collected during the project's implementation, ensuring optimised access and re-use of research data in alignment with the FAIR principles (Findable, Accessible, Interoperable and Reusable).

This version (V3) reflects the most up-to-date status of data management activities within the project, summarising the datasets effectively produced, shared, and preserved. It replaces the initial forecasts of previous DMP versions (V1 and V2) with verified information and links to the data repositories where the open datasets are stored.

All datasets generated within SEMPRES-BIO have been stored in secure institutional databases managed by each partner and in the project's shared Google Drive repository ("SEMPRES-BIO PROJECT") used as an internal collaborative platform. Data intended for open access have been published through the library section of the project website, peer-reviewed scientific publications, and Open Research Europe. In addition, selected datasets and related documentation have been deposited in open repositories such as Zenodo and are publicly accessible through persistent identifiers (DOIs).

Main categories of datasets generated within the project include:

- Project deliverables (technical and management reports).
- Articles published in Open Access scientific journals.
- Conference and workshop materials (abstracts, papers, posters).
- Meeting documentation (agendas, presentations, minutes).
- Multimedia materials (videos, infographics, promotional content).
- Communication materials hosted on the project website, social networks, and e-newsletters.
- Data generated in the Demonstration Sites (EBIEs) related to biomethane production efficiency, process parameters, and environmental performance indicators.

A summary table of the principal datasets is provided below, detailing the type, access conditions, format, and repository link:

**Table I. Summary of datasets generated within SEMPRES-BIO**

Dataset ID	Title	WP	Type of PID	Format	Repository / Link	Access Level	DOI / Licence
SB-WP2-D2.1	Experimental results from bench-scale testing used to guide the selection of reactor configurations and microbial cultures for the French and Spanish demo sites.	WP2	Other (Experimental)	XLSX	Internal (SEMPRES-BIO shared folder)	Restricted	N/A
SB-WP4-D4.1	Survey conducted by DBFZ among SEMPRES-BIO partners, three Horizon Europe projects, and external stakeholders on CO <sub>2</sub> valorisation practices in biomethane plants. Data used in Deliverable D4.1.	WP4	Technical / Tabular	Tabular PDF (with tables)	Project website - Library section <a href="https://sempre-bio.com/wp-content/uploads/2024/10/SB-WP4-D4.1-OpportunitiesValorizationCO2-PU.pdf">https://sempre-bio.com/wp-content/uploads/2024/10/SB-WP4-D4.1-OpportunitiesValorizationCO2-PU.pdf</a>	Open access	N/A

The SEMPRES-BIO consortium has ensured that all scientific publications resulting from the project are available in full open access, in compliance with Article 17 of Annex 5 of the Grant Agreement. In

line with Horizon Europe rules, no hybrid open-access publications or embargo periods have been applied, and only publication fees in full open-access venues were claimed as eligible costs.

The datasets and publications generated contribute to the long-term goal of enabling transparency, reproducibility, and wider re-use of research outcomes supporting the biomethane value chain. Data preservation beyond the end of the project will be ensured by CETAQUA and partner institutions for a minimum period of five years, following Horizon Europe recommendations.



### 3. FAIR Data

The SEMPRES-BIO project follows the FAIR guiding principles for scientific data management and stewardship (mandatory for defined open data) that are used all across the length of the project. As per the FAIR principles data should be:

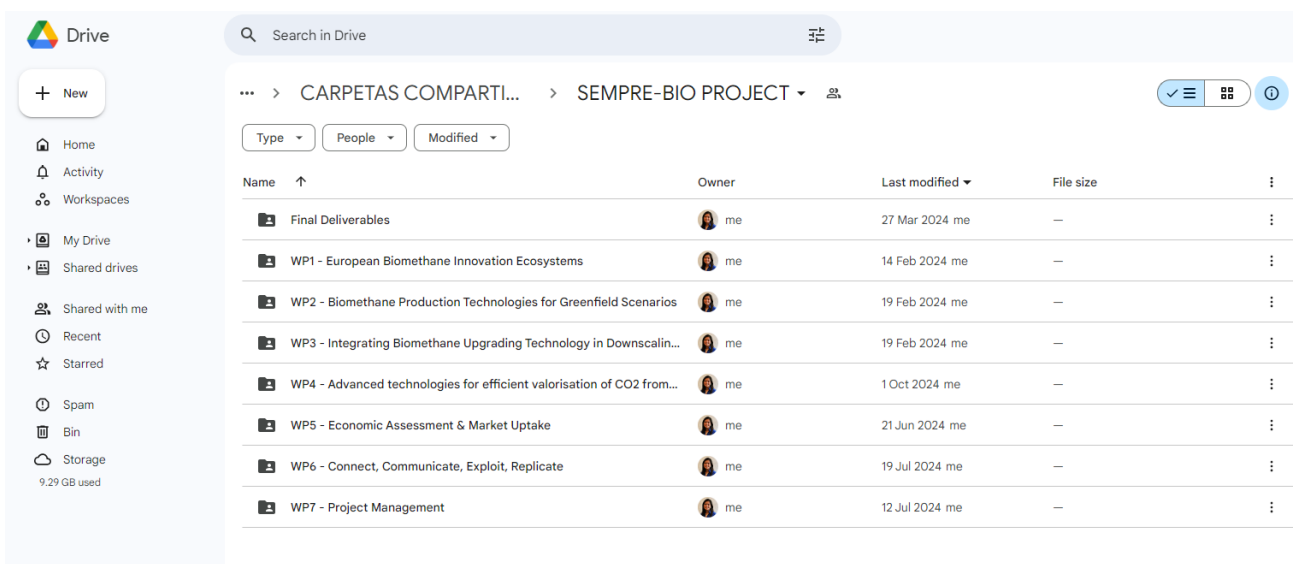
1. Findable.
2. Accessible.
3. Interoperable.
4. Reusable.

Access to data is kept as open as possible, but also as closed as necessary for commercial and exploitation purposes. Some key elements defined to ensure open access to data under this DMP are the following:

- Types of data/research outputs: reports, tools, videos, website, policy recommendations and scientific publications.
- Findability of data/research outputs: Digital Object Identifiers (DOI) have been assigned to SEMPRES-BIO publications, standardised metadata frameworks and ORCID identifiers are used to ensure data traceability and credit to individual researchers. Publications are accessible through Open repositories, and partners to a fully open research publishing model.
- Accessibility of data/research outputs: Full open access is provided to 14 out of 39 of SEMPRES-BIO deliverables, all publicly available to maximise dissemination and impact. The remaining 11 public deliverables will be made openly accessible upon finalisation and internal validation. Restricted deliverables (D2, D3, D6, D8, D10, D11, D12, D13, D15, D16, D18, D19, D20, D27) are registered as sensitives and their access is limited under the conditions of the GA.
- Interoperability of data/research outputs: Open-source fonts such as Barlow / Alatsi Fonts are used for reports ensuring consistent on-screen display and compatibility across platforms.
- Reusability of data/research outputs: Open datasets have been shared whenever possible to enhance reproducibility. Additional datasets will be made available as soon as the corresponding research activities are completed and validated. An analysis of sensitive data is performed to avoid any unintended disclosure or data leakage.
- Curation and storage/preservation costs: CETAQUA is responsible for the storage and preservation of data.

#### 3.1. Making data findable

The data management plan supports the effective collection and integration of SEMPRES-BIO project data. Project-internal storage, processing, and sharing of project data will mainly happen through the project's folder Google Drive "SEMPRES-BIO PROJECT", which is protected from unauthorised access and is managed by the project coordinator.



**Figure 1. Front-page of the Google Drive folder.**

All partners are invited to use it to exchange information and documents on project activities and any project related issues. A folder structure with major category divisions based on WPs already exists and partners can add more sub-folders as necessary or appropriate, but they must not remove any of the current folders or data. Access rights to the shared Google Drive are managed by the project coordinator. In case of technical limitations or access restrictions for specific partners, alternative exchange routes (e.g., via coordinator upload or shared email channels) are provided to ensure full data accessibility. Within the folder WP7 Project Management and available to all partners, there are the Grant Agreement, Consortium Agreement, minutes for the different project meetings and an updated contact list.

The transfer of project information by email is limited to a strict minimum. Public project data are mainly shared through the project website, social media, community, other dissemination channels and open repositories. It is ensured that public project data do not contain critical personal data or confidential information. Each data contributor, along with the designated reviewers, is responsible for preliminary checks to exclude any sensitive information before data submission. The final verification will be conducted by the Project Coordinator, in collaboration with the Data Manager (SINTEF), who will ensure all data shared publicly complies with GDPR and Horizon Europe data security standards.

In addition to project deliverables and datasets, several scientific publications have been made publicly accessible in open-access journals and conference proceedings, including those published in Chemical Engineering Journal, Bioresource Technology, and Waste Management Bulletin, as well as in the 17th International Conference on Greenhouse Gas Control Technologies (GHGT-17). These publications contribute to the project's commitment to transparency, knowledge dissemination, and FAIR data principles.

As of this version, 14 deliverables (36% of all project deliverables) have been made publicly available through the project website (<https://sempre-bio.com/>), ensuring open access and dissemination. The remaining 11 deliverables are scheduled for publication in the coming months and will also be released under open-access conditions once finalised and approved by the consortium. Access permissions for sensitive datasets remain restricted in compliance with the Grant Agreement (Articles 16 and 17), while metadata for these datasets have been made publicly available to facilitate discoverability without compromising confidentiality.

### 3.1.1. Data discoverability

---

Considering the FAIR data principles, public project data should be Findable by:

- Being assigned to a globally unique and persistent identifier.
- Containing enough metadata to fully interpret the data.
- Data and metadata being indexed in a searchable source.

By applying these principles data becomes Findable.

As of October 2025, a total of five scientific publications generated by the SEMPRES-BIO consortium have been published in open repositories and are easily discoverable through the project website's *Library* section, each assigned a Digital Object Identifier (DOI) to ensure long-term findability and citation.

Examples include:

- Biofilm mass transfer and thermodynamic constraints in biomethane production systems, *Chemical Engineering Journal*, DOI: 10.1016/j.cej.2024.156629.
- Carbon monoxide inhibition on acidogenic glucose fermentation and aceticlastic methanogenesis, *Bioresource Technology Journal*, DOI: 10.1016/j.biortech.2024.131076.
- From microbial heterogeneity to evolutionary insights: A strain-resolved metagenomic study of H<sub>2</sub>S-induced changes in anaerobic biofilms, *Chemical Engineering Journal*, DOI: 10.1016/j.cej.2024.149824.
- Modelling of Bio-Methanation, a Promising Route for GHG Emission Reduction, *SSRN Electronic Journal*, DOI: 10.2139/ssrn.5016466.
- Pilot-scale anaerobic digestion of on-farm agro-residues: Boosting biogas production and digestate quality with thermophilic post-digestion, *Waste Management Bulletin*, DOI: 10.1016/j.wmb.2025.100201.

The consortium will continue depositing future deliverables and scientific publications in trusted repositories to obtain additional DOIs and ensure long-term preservation.

### 3.1.2. Data identification mechanisms

---

All documents associated with the project will be identified with a unique, and persistent, identifier that will be given at the time of the submission process.

Deliverables: SB-WPx-Dx.x-Deliverable name-Dissemination level-vx.x

- Project acronym.
- Related Work Package number.
- Deliverable number and name.
- Dissemination level [Sensitive (SEN) or Public (PU)].
- Version number [vx.x].

**Example:** SB-WP1-D7.10-Project Management Plan-PU-v1.0

Other documents: SB-WPx-Name-Vx.x

- Project acronym.
- Optional: Related Work Package number.
- Title of the document.
- Optional: Version number [vx.x] / Date

**Example:** SB-CET minutes or SB-WP7-CET minutes-v1.1

### 3.1.3. Naming conventions

The project document naming conventions are as per the list below developed at the University of Edinburgh<sup>1</sup>:

1. Keep file names short, but meaningful.
2. Avoid unnecessary repetition and redundancy in file names and file paths.
3. Use capital letters to delimit words, not spaces or underscores.
4. When including a number in a file name, always give it as a two-digit number, i.e., 01-99, unless it is a year or another number with more than two digits.
5. If using a date in the file name, always state the date 'back to front', and use four-digit years, two-digit months and two-digit days: YYYYMMDD or YYYYMM or YYYY.
6. When including a personal name in a filename give the family name first followed by the initials.
7. Avoid using common words such as "draft" or "letter" at the start of file names, unless doing so will make it easier to retrieve the record.
8. Order the elements in a file name in the most appropriate way to retrieve the record.
9. The file names of records relating to recurring events should include the date and a description of the event, except where the inclusion of any of either of these elements would be incompatible with rule 2.
10. The file names of correspondence should include the name of the correspondent, an indication of the subject, the date of the correspondence and whether it is incoming or outgoing correspondence, except where the inclusion of any of these elements would be incompatible with rule 2.
11. The file name of an email attachment should include the name of the correspondent, an indication of the subject, the date of the correspondence, 'attach', and an indication of the number of attachments sent with the covering email, except where the inclusion of any of these elements would be incompatible with rule 2.
12. The version number of a record should be indicated in its file name by the inclusion of 'v' followed by the version number and, where applicable, 'DRAFT'.
13. Avoid using non-alphanumeric characters in file names.

### 3.1.4. Document versioning

Only documents created by the consortium will be versioned. To this end, document templates include a history panel (Table 1) for keeping track of the document versioning.

**Table 2. Document template extract showing the history panel.**

VERSION	DATE	AUTHOR	DESCRIPTION OF CHANGE

Project partners will identify different document versions by using a two-digit version number (vx.x). A document reviewed by another partner should be returned to the principal author by including the secondary author in the history panel. Only the principal author will change the version number at the document name.

<sup>1</sup> Gryzbowski, A., University of Edinburgh: (July 2007). Naming Conventions for Electronic Records. Retrieved from <https://www.ed.ac.uk/records-management/guidance/records/practical-guidance/naming-conventions> [accessed on April 2023]

### 3.1.5. Metadata standards

Basic metadata will be used to facilitate the efficient recall and retrieval of information by project partners. Additionally, it will contribute to facilitating the discoverability of the requested information. To this end, all documents related to the project have to include, in the second-page, an information panel with information about the project and the deliverable: deliverable number and title, WP number and leader, contributor(s), the author(s) and reviewer(s), date and dissemination level. The SEMPRE-BIO partnership will use templates and tables to display this information in the most homogenised way. For instance, Table 2 depicts the information panel of project documents.

**Table 3. Document template example showing the information panel on the front page.**

#### PROJECT INFORMATION

GRANT AGREEMENT NUMBER		101084297
PROJECT TITLE	SEcuring doMestic PRoduction of cost-Effective BIOmethane	
PROJECT ACRONYM	Sempre-Bio	
FUNDING SCHEME	HORIZON-IA	
START DATE OF THE PROJECT	1 November 2022	
DURATION	42 months	
CALL IDENTIFIER	HORIZON-CL5-2021-D3-03-16	
PROJECT WEBSITE	<a href="https://sempre-bio.com/">https://sempre-bio.com/</a>	

#### DELIVERABLE INFORMATION

DELIVERABLE N°	
DELIVERABLE TITLE	
WP NO.	
WP LEADER	
CONTRIBUTING PARTNERS	
AUTHORS	
REVIEWERS	
CONTRACTUAL DEADLINE	
DELIVERY DATE TO EC	
DISSEMINATION LEVEL	

When depositing a publication in a repository, the required metadata will be entered, especially the funding programme (Horizon Europe), the associated European Research Council (ERC) grant number and the grant acronym, so that the publications can be linked to this project.

To further support data discoverability and interoperability, SEMPRE-BIO will adopt standardised metadata schemes. The DataCite schema will be used for publications and datasets, ensuring consistent citation and accessibility, while DCAT (Data Catalog Vocabulary) will structure the project's data catalogues. Together, these approaches will standardise documentation, facilitating data access, integration, and reuse across different platforms.

## 3.2. Making data accessible

### 3.2.1. Data availability

Project data, dissemination materials and deliverables that are declared as public are made available through the project website, social media channels, community and/or open repositories. Furthermore, the pseudonymized data sets are exploited through the creation of tables and infographics that are included as part of the dissemination activities of the project.

As described in Section 3.1., 14 SEMPRES-BIO deliverables have already been made publicly available through the project website, while the remaining ones will be released under open-access conditions once finalised. This section focuses on the accessibility of the project's open dataset and related research outputs.

In line with the FAIR principles, the main open dataset produced within SEMPRES-BIO is Deliverable D4.1 – "Opportunities for the Valorisation of CO<sub>2</sub>" (WP4). This deliverable, publicly accessible through the project website (<https://sempre-bio.com>), includes structured and reusable data describing CO<sub>2</sub> streams, process configurations, and potential valorisation pathways across the demonstration sites. Although published as a technical report (PDF), D4.1 contains quantitative datasets in tabular form that can be reused for techno-economic or environmental assessments related to carbon capture and utilisation. Metadata such as title, authorship, dissemination level, and publication date are embedded in the document, ensuring traceability and findability.

As detailed in Section 3.1.1, peer-reviewed publications in open-access journals (e.g., Chemical Engineering Journal, Bioresource Technology, Waste Management Bulletin) complement the dissemination of SEMPRES-BIO results and ensure the broader accessibility of project findings.

### 3.2.2. Data access

In order to maximise the impact of SEMPRES-BIO research data, the results are shared within and beyond the consortium. Selected data and results are made accessible to the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through open access data repositories.

SEMPRES-BIO project results and deliverables are first stored and organised in the project's folder Google Drive, on the personal computers of project staff members, on the institutional secure server and on the project website. All data are made available for verification and re-use, unless the task leader can justify why data cannot be made openly accessible.

When applicable, Creative Commons licences (e.g., CC BY 4.0) will be applied to future open-access deliverables and datasets to ensure proper citation and re-use. Currently, publicly available deliverables are accessible through the project website under the dissemination level "PU" (Public), in compliance with Horizon Europe open-access requirements.

Project outcomes and public deliverables are accessible primarily through the SEMPRES-BIO project website (<https://sempre-bio.com>), which serves as the main platform for dissemination and open access. Data and documents intended for internal use only are stored in the consortium's shared Google Drive repository and institutional servers.

In the future, public datasets may also be deposited in trusted open repositories to ensure long-term preservation and DOI assignment.

Access to restricted or confidential information can be granted upon justified request to the project coordinator or data owner. When necessary, appropriate Intellectual Property Rights (IPR) procedures – such as Non-Disclosure Agreements (NDAs) – will be applied to protect sensitive or proprietary data.

### 3.3. Making data interoperable

The SEMPRES-BIO consortium aims to ensure that all datasets and research outputs are stored and managed in formats that facilitate interoperability and long-term reuse by external users and stakeholders.

Although no datasets have yet been deposited in open repositories, the data produced within the project (e.g., deliverables, tabular results, and reports) are already stored in widely readable formats such as PDF, Excel (.xlsx), and CSV, which can be easily converted into open, machine-readable formats for future use.

The project partners will pay reasonable efforts towards storing all data in the appropriate format that will make data re-usable by external parties that might be interested in exploiting the data generated during the project. Partners will observe OpenAIRE guidelines for online interoperability, including OpenAIRE Guidelines for Literature Repositories, OpenAIRE Guidelines for Data Archives, OpenAIRE Guidelines for Current Research Information Systems (CRIS) Managers based on CERIF. These guidelines can be found at: <https://guidelines.openaire.eu/en/latest/>

The default language is British English, and the vocabulary used follows common terminology within the research and manufacturing sectors, addressing the specific target audience of Small and medium-sized enterprises (SMEs), large companies, and their ecosystems. Vocabulary will not constitute a barrier for data interoperability and reuse. A project glossary has been set up and will be incremented as time goes.

The following table describes a variety of file formats for different disciplines that are either recommended or acceptable:

**Table 4. File formats for different disciplines that are either recommended or acceptable**

Type of data	Recommended formats	Acceptable formats
Tabular data	CSV (.csv), Tab-delimited (.tab)	Excel (.xls/.xlsx)
Geospatial data	ESRI Shapefile (.shp), GeoTIFF (.tif)	Geography Markup Language (.gml)
Textual data	PDF/A, RTF (.rtf)	DOCX (.docx), HTML (.html)
Image data	TIFF (.tif)	JPEG (.jpg), PNG (.png)
Video data	MPEG-4 (.mp4)	AVCHD video (.avchd)
Documentation and scripts	PDF/A, XHTML	DOCX (.docx), XML (.xml)

This selection is based on recommendations for data preservation from the OpenAIRE guidelines for Horizon Europe projects. For a full list of formats by data type, refer to the [OpenAIRE Guidelines for Data Archives](#)

### 3.4. Increase data re-use

To enhance data re-use, the SEMPRES-BIO consortium ensures that all public datasets and deliverables are made available under clear and transparent conditions.

Currently, the main open dataset produced within the project corresponds to Deliverable D4.1 – “Opportunities for the Valorisation of CO<sub>2</sub>” (WP4), which is publicly accessible through the project website and may be deposited in a trusted open repository (e.g., Zenodo or OpenAIRE) at a later stage.

Appropriate Licensing will be used to protect the ownership of the datasets. Depending on the need for attribution and reuse conditions, licences such as the Open Database License (ODbL) or the Public

Domain Dedication and License (PDDL) will be considered for the parts of datasets for which the decision of making that part public has been made by the Consortium.

For any future datasets deposited in a public data repository, access will be open and unlimited. Restrictions on-use policy are applied for all protected data, whose re-use will be limited within the project partners. Specific restrictions may include the following:

- Attribution: requires users of the dataset to give appropriate credit, provide a link to the licence, and indicate if changes were made
- Limited to Non-Commercial use: prohibits the use of the dataset for commercial purposes by others.
- Share under the same terms: requires the others to use the same licence as the original on all derivative works based on the original data.

An internal process of Quality evaluation is activated throughout the entire project duration to assess both project data /products and project process. Internal peer review is performed for the project deliverables to guarantee the deliverable is developed with a high level of quality(see the D7.11 Project Management Plan).

## 4. Allocation of resources

Data management in SEMPRE-BIO is coordinated under the WP7. CETAQUA, as project coordinator, is responsible for ensuring compliance with FAIR data management principles and maintaining the Data Management Plan. A portion of the WP7 budget and person-months has been allocated to these activities, covering tasks such as DMP updates, partner support on data handling, and supervision of open-access deliverables.

To date, the project coordinator has overseen the implementation of FAIR data management practices and ensured that all public deliverables and publications comply with the Horizon Europe open-access requirements.

Resources for long-term preservation and the associated costs will be defined and agreed upon by the consortium before the end of the project, within the Project Steering Board (PSB).

The consortium will evaluate options for maintaining accessibility beyond the project duration – including preservation of public deliverables and datasets via the project website and, if feasible, through trusted open repositories to ensure long-term visibility and reusability of SEMPRE-BIO results.

## 5. Data security

All project deliverables and data have been stored and shared in the folder name "SEMPRE-BIO PROJECT" in Google Drive restricted to the project consortium. As an initial step, only the Consortium Partners will have access to the cloud storage where dataset and metadata are filed. Following scientific publications and articles, the dataset deliverables and the final demonstrator research results will be shared through open repositories and other databases to promote the data making FAIR.

## 6. General Data Protection Regulation (GDPR)

The SEMPRES-BIO project adheres to the General Data Protection Regulation (GDPR), officially known as Regulation (EU) 2016/679, to ensure the protection of personal data collected, processed, and shared throughout the project lifecycle. This compliance ensures that all personal data handling respects the rights of individuals while enabling legitimate project operations.

The GDPR is a comprehensive legal framework that governs the processing of personal data within the European Union (EU) and the European Economic Area (EEA). Enforced since May 25, 2018, the GDPR aims to harmonize data privacy laws across Europe, protect and empower all EU citizens' data privacy, and reshape the way organizations across the region approach data privacy.

A summary of this legal framework:

### Key Objectives of the GDPR:

1. **Protection of Personal Data:** The GDPR safeguards individuals' fundamental rights and freedoms concerning the processing of personal data, ensuring that personal information is handled with care and respect for privacy.

### Core Principles:

The GDPR outlines several key principles that organizations must adhere to when processing personal data:

- **Lawfulness, Fairness, and Transparency:** Personal data must be processed legally, transparently, and fairly.
- **Purpose Limitation:** Data should be collected for specified, explicit, and legitimate purposes related to project objectives.
- **Data Minimization:** Data collected must be adequate, relevant, and limited to what is necessary.
- **Accuracy:** Personal data must be accurate and, where necessary, kept up-to-date.
- **Storage Limitation:** Data should be kept in a form that permits identification of data subjects for no longer than is necessary.
- **Integrity and Confidentiality:** Personal data must be processed securely to prevent unauthorized access, disclosure, or loss.

### Data Categories and Processing Activities

The SEMPRES-BIO project will process a limited amount of personal data, strictly necessary for project implementation. The following categories of personal data may be processed:

- **Contact information:** Includes names, institutional affiliations, and email addresses of consortium members and stakeholders involved in project-related activities.
- **Survey, interview, or focus group data:** If applicable, anonymized or pseudonymized responses from research participants, collected to assess stakeholder perspectives on biomethane production and related topics.
- **Processing Activities:** Data collection, transfer, storage, and analysis will be performed in compliance with GDPR. Personal data will only be collected when necessary for project execution (e.g., stakeholder engagement, surveys, interviews) and will be securely stored on Google Drive (restricted access) and institutional databases of project partners. Data may be shared internally among project partners for research purposes, and if data is published, it will be either anonymized or pseudonymized to ensure compliance with GDPR.

### Rights of Data Subjects:

The GDPR enhances existing rights and introduces new rights for individuals, including:

- **Right of Access:** Individuals can access their data.

- Right to Rectification: Individuals can request the correction of inaccurate personal data.
- Right to Erasure ("Right to be Forgotten"): Under certain conditions, individuals can request the deletion of their personal data.
- Right to Restriction of Processing: Individuals can request the limitation of how their data is processed under specific circumstances.
- Right to Data Portability: Individuals can transfer their data between service providers.
- Right to Object: Individuals can object to the processing of their personal data in certain situations, including direct marketing.

### **Obligations for Data Controllers and Processors:**

Organizations that handle personal data have specific responsibilities under the GDPR:

- Implement data protection measures by design and by default.
- Conduct Data Protection Impact Assessments (DPIAs) when necessary.
- Appoint Data Protection Officers (DPOs) where required.
- In the event of a personal data breach, organizations must notify the relevant supervisory authority within 72 hours when risks are significant.

In summary, the GDPR establishes a robust framework for data protection within the EU and EEA, emphasizing transparency, accountability, and the empowerment of individuals regarding their personal data. Organizations involved in processing personal data must ensure compliance with the regulation to uphold individuals' rights and avoid substantial penalties.

## 7. Ethical Aspects

The deliverable D7.8 Research Ethics and Data Protection monitoring Report V1 aims at ensuring that ethical requirements are met for all research undertaken in the project, in compliance with Horizon Europe standards. All partners will assure that the EU standards regarding ethics and data management are fulfilled in compliance with the ethical principles (see Article 14 and Annex 5 as set out in the GA). In addition, SEMPRE-BIO partners have to comply with:

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data and on the free movement of such data (GDPR).
2. Ethical principles (including the highest standards of research integrity as set out, for instance, in the European Code of Conduct for Research Integrity and including avoiding fabrication, falsification, plagiarism or other research misconduct).

A second and final version (V2) of Deliverable D7.8 will be prepared at the end of the project to summarise the monitoring results of research ethics and data protection compliance across all project activities, ensuring consistency with the updated GDPR framework (Section 6 of this DMP).

## 8. Other issues

Researchers will ensure that all aspects of their research data management are thoroughly planned and documented, which will contribute to improving the quality, transparency, and accessibility of their research. To this end, as well as European Commission policies on open data management, Project Partners must also adhere to their own institutional policies and procedures for data management.

## 9. History of Changes

Table 5. History of changes.

Version Number and date	Change and justification	Section of D7.5
2.0 07/10/2024	Updated the number and name version of deliverable to correspond to the new version of the document for the previous deliverable.	Introduction (Section 1)
2.0 07/10/2024	Updated with the name of the new beneficiary to the project.	Consortium partners
2.0 07/10/2024	Removed Basecamp, as it is no longer available. Google Drive is now the platform to share documents, replacing Basecamp.	Data Summary (Section 2) FAIR Data (Section 3.1, 3.2.2) Data Security (Section 5)
2.0 07/10/2024	Replaced the Basecamp figure with Google Drive folder, as this platform is no longer used.	FAIR Data (Section 3.1)
2.0 07/10/2024	Updated the identified mechanism of the other documents, making the related Work Package number optional to the document name.	FAIR Data (Section 3.1.2)
2.0 08/10/2024	Updated the name of Project Steering Board (PSB)	Allocation of resources (Section 4)
2.1 28/10/2024	Defined roles for initial and final data checks	Making data findable (Section 3.1)
2.1 28/10/2024	Specified that Data and metadata applied for indexing in a searchable source.	Data discoverability (Section 3.1.1)
2.1 28/10/2024	Added the standardised metadata schemes.	Metadata standards (Section 3.1.5)
2.1 28/10/2024	Included recommended file formats for different disciplines.	Making data interoperable (Section 3.3)
2.1 28/10/2024	Updated licensing information to emphasize appropriate licensing options over Creative Commons.	Increase data re-use (Section 3.4)
2.3 13/12/2024	Added a new chapter on General Data Protection Regulation (GDPR)	General Data Protection Regulation (GDPR) (Section 6)
3.0 28/10/2025	Updated the deliverable number and version to D7.6 V3. Integrated feedback from previous periodic reviews.	Cover page, document log

3.0 28/10/2025	Rewritten "Data Summary" to reflect the current datasets produced, published, and preserved, including table of datasets and repository links.	Data Summary (Section 2)
3.0 28/10/2025	Updated "FAIR Data" sections with verified information, DOIs of publications, and statistics of open deliverables (14/39)	FAIR Data Section 3, 3.1 and 3.1.1
3.0 28/10/2025	Added details on dataset accessibility (D4.1), licensing and data sharing mechanisms.	Making data accessible Section 3.2
3.0 28/10/2025	Refined interoperability section to include current file formats and OpenAIRE compliance statement.	Making data interoperable Section 3.3
3.0 28/10/2025	Revised "Increase data re-use" to include references to D4.1 dataset, licensing options, and internal quality control.	Increase data re-use Section 3.4
3.0 28/10/2025	Updated "Allocation of resources" to reflect current implementation status and future preservation plans.	Allocation of resources Section 4
3.0 28/10/2025	Expanded "Ethical Aspects" with reference to D7.8 V2 and alignment with GDPR compliance.	Ethical Aspects Section 7