D1.2 – Data Management Plan

Deliverable:	Data Management Plan
Author(s):	Claudio Carbone and Alessandro Agostini (ENEA)
Version:	Final
Quality review:	Stefano Proietti and Giorgia Galvini (ISINNOVA)
Date:	01/03/2023
Dissemination level:	Public (PU)
Grant Agreement N°:	101084200
Starting Date:	01-10-2022
Duration:	54 months

Coordinator:	Stefano PROIETTI, ISINNOVA
Tel:	0039 06. 32.12. 655
Fax:	0039 06. 32.13. 049
E-mail:	sproietti@isinnova.org

Co-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.

Co-funded by the European Union

CONTENTS

Executive Summary 4
1. BIOMETHAVERSE in a nutshell 5
2. Introduction 6
2.1. Deliverable overview and structure 6
2.2. Administrative information 7
2.3. Data management, responsibilities & resources 7
3. Data Management Plan 9
3.1. Data summary 9
3.2. Data acquisition and formats 10
3.3. Data overview per Work Package 11
3.4. Generated research data 14
3.4.1. FAIR principles 15
3.4.2. Metadata produced from the project 16
3.4.3. Data processing & reuse of existing data 17
3.5. Intellectual Properties Rights (IPR) 17
4. Data storage and management 17
4.1. Data archiving and preservation 17
4.2. Which data need to be preserved? 18
4.3. What is potentially useful to others? 18
4.4. What has scientific, cultural or historical value? 18
4.5. What legally must be destroyed? 18
4.6. Data storage during the process 18
4.7. Data sharing and dissemination 19
5. Data Ethics and Security 19
5.1. Ethics – Personal Data 19
5.1.1. Dissemination, communication & exploitation activities data usage processes 19
5.1.2. BIOMETHAVERSE mailing list 19
5.1.3. Meeting related material 20
5.1.4. Workshops/Conferences and Training sessions 20
5.1.5. Newsletters, social media, and other dissemination material 20
5.1.6. Usage of cookies (in BIOMETHAVERSE website) 20
5.2. Informed Consent Procedures 20
5.3. Legal Aspects 20

Co-funded by the European Union

5.3.1. Embargo periods 21
5.3.2. Sensitive data 21
5.3.3. Data leaks 21
5.4. Ethical aspects 21
5.5. Security 21

List of Figures

Figure 1 BIOMETHAVERSE countries and partners 5
Figure 2 Screenshot of the SharePoint Microsoft Teams – storage location of the corresponding WP data 9
Figure 3 Screenshot of the repository Zenodo 9
Figure 4 BIOMETHAVERSE IPR process. 17

List of Tables

Table 1 Adherence to BIOMETHAVERSE deliverable and task description 6
Table 2 List of personnel for data management 8
Table 3 List of leads of each work package 10
Table 4 Insight into data information/generation from each WP 11
Table 5 Data ownership attributes 14

Co-funded by the European Union

Executive Summary

Deliverable 1.2 is the Data Management Plan (DMP) of the BIOMETHAVERSE project and it is intended as a living document which outlines how the data produced within the BIOMETHAVERSE project will be managed to make them findable, accessible, interoperable, and reusable.

D1.2 is addressing WP1 'Project Coordination and Management' and the work under Task 1.2 'Management of data and other research output'.

This document includes information covering all the areas below:

- The guiding principles for data management in general within the project
- The legal framework constituted by the General Data Protection Directive (GDPR)
- Data Summary: Overview of what data will be generated, gathered, and processed during the course of the project (including, where applicable, personal data)
- Data storage: How data will be stored and processed according to the Findable, Accessible, Interoperable and Reusable (FAIR) Data Management principles, making data: findable, accessible, interoperable, and reusable.
- Metadata: Why Metadata is important and how metadata will be made available from different activities in the project.
- Data Ethics and Security: How the Consortium intends to keep the data secure and how that can be shared with the public.

The project runs for a total duration of 54 months and during this period, the data generated must be handled in a responsible and useful manner and this is exactly the intention of the DMP. To facilitate data exchange, a SharePoint (Microsoft365 TEAM) has already been created and this provides safe and secure access to all project participants. Data security is carefully taken into consideration, and there are measures in place (if needed) to cordon off certain areas of SharePoint in case sensitive information is to be stored or shared. Also, access to SharePoint is only given to personnel who are working on specific tasks within the project. All other public data will be available to all participants.

It is the intention of the project to maximize its impact by sharing information as publicly as possible. In this regard, all scientific publications will be made open access and except for any sensitive information, all other information will be disseminated or published in an open-access platform.

The DMP currently is in its first version, and it is intended to provide maximum clarity and information to all relevant stakeholders and partners. This initial version of the DMP defines the general policy and approach to data management adopted in BIOMETHAVERSE and it handles data management-related issues at the administrative and technical level.

As the data identification and collection activities are still ongoing, the initial DMP can currently only provide an incomplete picture of the datasets that are needed in the different demonstration cases and the BIOMEHTAVERSE project in general. Final adjustments and refinements will be made at the end of the project (M50, D1.3) according to the Grant Agreement (GA), to present the actual research data generated during the project and include updated instructions for how to access open data.

Co-funded by the European Union

1. BIOMETHAVERSE in a nutshell

BIOMETHAVERSE (Demonstrating and Connecting Production Innovations in the **BIOMETHA**ne uni**VERSE**) aims to diversify the technology basis for biomethane production in Europe, increase its cost-effectiveness, contribute to the uptake of biomethane technologies, and support the priorities of the SET Plan Action 8.

To meet these goals, **five innovative biomethane production pathways** will be demonstrated in five European countries: France, Greece, Italy, Sweden, and Ukraine.

The five selected demonstrators go beyond the state of the art and thus beyond technologies already implemented at a commercial scale and rely on:

- In-situ and Ex-Situ ElectroMethanoGenesis (EMG): Electricity enhanced biomethane production (by ENGIE, France);
- Ex-situ Thermochemical/catalytic Methanation (ETM): Thermochemical/catalytic upgrading of biogas using hydrogen (by BLAG, Greece);
- Ex-Situ Biological Methanation (EBM): Biological upgrading of biogas using hydrogen, including feed-stock pre-treatment via ozonolysis (by CAP, Italy);
- Ex-Situ Syngas Biological methanation (ESB): Biological methanation of syngas from thermal gasification (by RISE, Sweden);
- In-situ Biological Methanation (IBM): Hydrogen integration in the AD reactor (by MHP, Ukraine).

The project's objectives will be achieved through the implementation and consolidation of the following founding pillars:

- Demonstration of Innovative Biomethane Pathways;
- Assessment and Optimisation of Innovative Biomethane Pathways;
- Replicability, Planning Decisions, Market Penetration, and Policy Dimension;
- Dissemination, Exploitation & Communication.

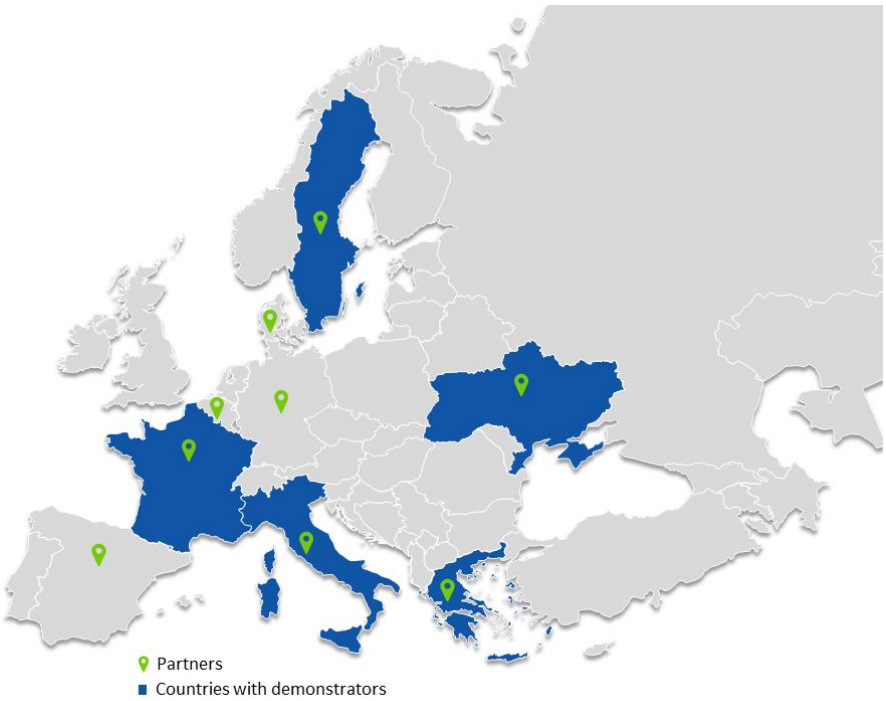

Figure 1 BIOMETHAVERSE countries and partners

Co-funded by the European Union

2. Introduction

The purpose of this section is to map BIOMETHAVERSE commitments, both within the formal Deliverable and Task description against the project's respective outputs and work performed.

Table 1 Adherence to BIOMETHAVERSE deliverable and task description

BIOMETHAVERSE task		Document chapter(s)	Justification
Task 1.2 Management of data and other research outputs	Task 1.2 Management of data and other research outputs (Lead: ENEA; participants: ISINNOVA) (M1-54) This task will treat knowledge and data management according with the HE rules regarding the open access of data and other research outputs, and it will produce a Data Management Plan (DMP), revised towards the end of the project, to clearly: • define the strategy for the management of data and other research outputs (including IPR aspects); • establish clear accountability for data collection, storage and management, including treatment of personal information; The DMP will include sections on: 1) Data Types, Formats, Standards and Collection Methods, 2) Access, Data Sharing and Reuse, 3) Resourcing, 4) Deposit and Long-Term Preservation, and 7) Short-Term Storage and Data Management. ENEA will set-up and maintain a Zenodo repository linked to OpenAire for open access data management.	Section 2 Section 3 Section 4 Section 5	Section 2 mainly provides administrative information and responsibility description on BIOMETHAVERSE data management. Section 3 presents data management specific to the project and how metadata needs to be handled and also IPR & innovation. Section 4 provides information on data archiving, storage and preservation which is critical as the project progresses. Section 5 talks about data security, legal and ethical aspects which are also crucial when vast amounts of data are to be handled.
BIOMETHAVERSE deliverable			
D1.2 Data Management Plan: The data generated from several tasks from different work packages need to be managed in a responsible manner. A detailed report on how this vast amount of data will be handled is described in this report.			

2.1. Deliverable overview and structure

The deliverable is organised in the following structure.

Section 2 provides administrative information, responsibilities for all involved project partners and resource information for storing and maintaining data.

Section 3 is about the DMP in general and with a focus on:

- Data Summary,
- Data acquisition formats,

Co-funded by the European Union

- Data overview per WP,
- How to handle generated research data, the creation of metadata & Intellectual Property Rights (IPR) involved. The IPR will be handled according to the text provided here and as per details given in the Consortium Agreement (CA).

Section 4 provides information on:

- Data archiving and preservation,
- Data storage aspects during research progress and
- Data sharing and dissemination during and after project phase.

This will help lay down protocols in place for data storage.

Section 5 deals with data ethics and security within which topics on ethics involved with personal data collection, informed consent procedures, legal and ethical aspects involving data, and data security will be touched upon.

Section 6 provides a summary of the entire report.

2.2. Administrative information

The project BIOMETHAVERSE under GA # 101084200 is coordinated by ISINNOVA with the main Project Coordinator (PC) being Stefano Proietti. Their respective contact information is as follows:

- Stefano Proietti – sproietti@isinnova.org

For data dissemination, communication, and exploitation, ISINNOVA has the lead as well.

2.3. Data management, responsibilities & resources

All data generated within the project, will be managed responsibly by ENEA and all information will be available for use by the project partners, subject to terms and conditions mentioned in the CA.

ISINNOVA, as project coordinator, has set up a TEAM with the Microsoft365 environment, which includes a SharePoint for collaboration and files/data exchange. ENEA has set up a ZENODO repository (https://zenodo.org/communities/biomethaverse) to guarantee the access to public data, deliverables, documents and scientific papers, enable their identification with a DoI, and guarantee their future accessibility after the project end.

Key personnel in the project – the PC Stefano Proietti and DMP task leader Claudio Carbone will act as data stewards during the timeline of the project and will be responsible for documenting and managing the data during this period.

Additionally, the main leads from every partner will be responsible for the following:

- Data management at their respective organisations/institutes/local consortium.
- Data transmission from their respective organisations/institutes/local consortium to the SharePoint.
- Ensuring no sensitive data flows out beyond the personnel working on the project.
- Ensuring no data is made public prior to joint consultation with the Consortium.
- Reporting to the PC in case they come across any issues related to data management and security.

The list of the main leads from the partners is mentioned in Table 2.

Co-funded by the European Union

Table 2 List of personnel for data management

Partner	Lead1	Lead2
Istituto di Studi per l'Integrazione dei Sistemi (ISINNOVA)	Stefano Proietti	Giorgia Galvini
European Biogas Association (EBA)	Mieke Decorte	Gabriella Papa
Agenzia Nazionale per le Nuove Tecnologie, l'Energia e lo Sviluppo Economico Sostenibile (ENEA)	Alessandro Agostini	Claudio Carbone
Bioenergy Association of Ukraine (UABIO)	Georgiy Geletukha	
Bioaerio Lagada (BLAG)	Themistoklis Sfetsas	
Ethniko Kentro Erevnas Kai Technologikis Anaptyxis (CERTH)	Dimitris Kourkoumpas	
RISE Research Institutes of Sweden (RISE)	Emelie Ljung	
Cortus Energy (CORTUS)	Staffan Hellsén	
Wartsila Sweden (WARTSILA)	Lars- Evert Karlsson	
ENGIE (ENGIE)	Gaspard BOUTEAU	
AERIS Tecnologias Ambientales (AERIS)	Óscar J. Prado Rubianes	
Acondicionamento Tarrasense Associacion (LEITAT)	Daniele Molognoni	
Danmarks Tekniske Universitet (DTU)	Yifeng Zhang	
Friedrich-Alexander-Universitaet (FAU)	Katharina Herkendell	
CAP Holding (CAP)	Maria Rosaria Scoppettuolo	
Energigas Sverige Service (SGA)	Linus Klackenberg	
PrJSC "MHP Eko Energy" (MHP)	Oleksandr Dombrovskiy	
Politecnico di Milano (POLIMI)	Francesca Malpei	
Consorzio Italiano Compostatori (CIC)	Alberto Confalonieri	
Società Italiana Acetilene e Derivati (SIAD)	Giorgio Bissolotti	
Deutsches Biomasseforschungszentrum gemeinnützige (DBFZ)	Jörg Kretzschmar	
Ellman Engineering (EE)	Reik Ellmann	

Every partner must ensure that data used during the project, for WPs, for tasks, for sub tasks and any other purpose is stored and made available later, if not protected for confidentiality or IPR. This might be needed either for scientific publications or post project dissemination or for creating metadata, all of which might happen after the project has ended. The respective lead or substitute lead from every partner must also ensure that all data generated at their premises is stored in a safe and secure manner and minimise the chances for data loss.

Resources needed for the SharePoint and data management on the SharePoint will be covered by the project management budget in the proposal. At this point of time, no additional costs are foreseen for the activities listed out in the project DMP and later version of the DMP are not expected to identify extra costs needed to store and maintain data.

The DMP is a living document, with the first version being published in February 2023 (M5) of the project.

Co-funded by the European Union

3. Data Management Plan

3.1. Data summary

A significant amount of data is expected to be generated along the course of the project and this amount of data needs to be managed in a responsible manner. The exact amount of data will be evaluated at the end of each reporting period.

The data arising or contributing from/to a respective WP must be stored under the WP sub-area on the SharePoint. It is the responsibility of the WP leader to ensure that the corresponding WP on the SharePoint is maintained with regards to data management. The leads from respective partners, especially the demonstrators' leaders, must also actively contribute to the data in respective WP's they are involved in by cooperating closely with the WP leader.
For example, data from WP3 will be directly stored under 'WP3' sub-area on the SharePoint. This will enable all partners involved in WP3 to easily access the data.

Figure 2 shows a screenshot of the SharePoint Microsoft Team, while Figure 3 shows a screenshot of the repository Zenodo.

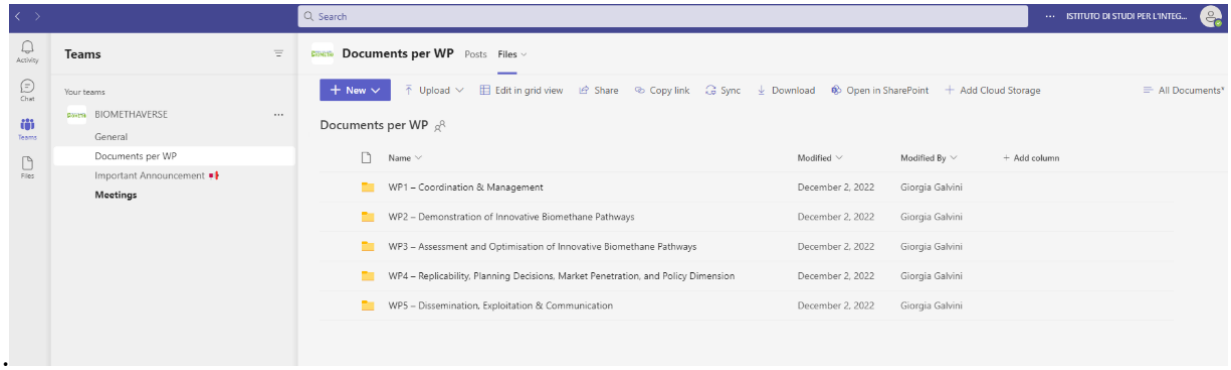

Figure 2 Screenshot of the SharePoint Microsoft Teams – storage location of the corresponding WP data

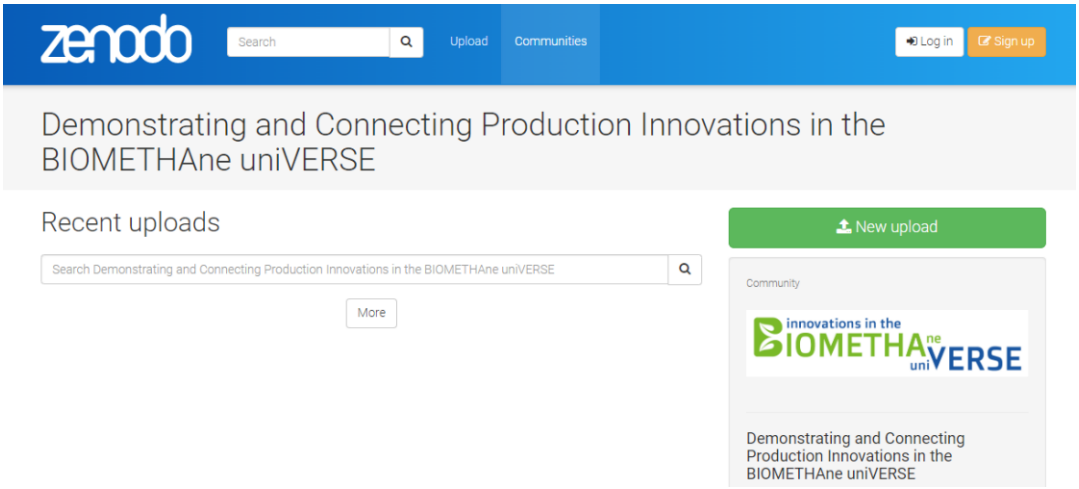

Figure 3 Screenshot of the repository Zenodo

Table 3 shows the partners who are leading the respective WPs.

Co-funded by the European Union

Table 3 List of leads of each work package

Work Package	Lead
WP1 – Coordination & Management	ISINNOVA
WP2 - Demonstration of Innovative Biomethane Pathways	EBA
WP3 - Assessment and Optimisation of Innovative Biomethane Pathways	ENEA
WP4 - Replicability, Planning Decisions, Market Penetration, and Policy Dimension	EBA
WP5 - Dissemination, Exploitation & Communication	ISINNOVA

Each WP leader has the explicit right to decide on which data needs to go under each WP sub-area and can also decide the accessibility rights for the data. The coordination of the respective WPs is left to the WP leader and together with the task leaders and demonstrators' leaders, they can decide on the data content and other data related management activities. The PC will intervene only when the handling of the data is not being carried out as per the guidelines laid out.

The purpose of data generation and data preservation is to demonstrate innovative biomethane production as energy carrier and a fuel through the five different production pathways at four different levels. The first is to make the demonstrator fully operational and increase their cost-effectiveness by fine-tuning different technological solutions; the second is to carry out a Life Cycle Sustainability Assessment (LCSA) and a Techno Economic Analysis (TEA) on such systems; the third is to ensure replicability and upscaling and the final one is to ensure market penetration.

All open data and publications produced in BIOMETHAVERSE will be identifiable and locatable by means of a persistent Uniform Resource Locator (URI) and, when possible, it will be assigned a Digital Object Identifier (DOI) in order to make content easily and uniquely citable.

Open results that are deposited in the BIOMETHAVERSE default Open Access repository (Zenodo) will be assigned a DOI automatically. Open results that are deposited in institutional repositories, repositories of scientific publishers or other data and research repositories will be at least indefinable by a persistent URI. If the institution is a DOI registrant that has an agreement with a DOI registration agency, a DOI will be assigned, too.

Whether scientific publications will be assigned a unique identifier like DOI, Publisher Item Identifier (PII), International Standard Serial Number (ISSN), etc. depends on the open access strategy) chosen by the editors and thus also on the respective scientific publisher and the chosen research repository. Zenodo (http://help.zenodo.org/features/) is for example, one of the open data repository repositories that can generate DOIs for research results.

3.2. Data acquisition and formats

Data will be acquired through various means in the project. Some of the data acquisition methods include but not limited to the following:

- For experimental data - via test equipment, measurement devices
- For simulation data – via software tools, data processing and analysis tools
- For statistical data – via data processing tools, google analytics, social media
- For personal data – google analytics, email.

Generated research data will be in (but not limited to) the following formats:

- docx
- xlsx
- pptx

Co-funded by the European Union

- pdf
- csv
- txt

Besides the above common formats, images/photos in the format of .png, .jpeg or any other relevant format, software codes and algorithms (in the respective formats of the software tools used), LCSA & TEA data in their respective formats, will also be generated.

Wherever applicable and possible, provisions will be provided for data formats to be migrated when new technologies/tools become available and are proved robust enough to not only ensure digital continuity but also availability of data.

On the TEAMS SharePoint, the following applies to the use of special characters in the file or folder name:

- One cannot use tilde (~), number sign (#), percent (%), Ampersand (&), Asterix (*), Braces ({}), backslash (\), colon (:), angle brackets (<>), question mark (?), slash (/), plus sign (+), pipe (|), quotation mark (“).
- One cannot use the period character consecutively in the middle of a file name.
- One cannot use period character at the end of a file name.
- One cannot start a file name by using the period character.
- Use of an underscore character (_) at the beginning of a file name will result in a hidden file.

In the BIOMETHAVERSE TEAMS SharePoint under WP5 – Dissemination, Exploitation & Communication - a series of documents/reports templates have been created to ensure a consistent approach for all BIOMETHAVERSE data and their versions and provide for the needed alignment to the project identity.

3.3. Data overview per Work Package

This section provides a data overview per WP, identifies the possible data sources and specific information of the generated data (Table 4).

Table 4 Insight into data information/generation from each WP

Work Package	Type of data information	Description
WP1	What are the data sources and where do they come from?	This is a project management WP, no specific scientific data will be generated from this WP. Instead, data from all other WPs will be used here for management purposes.
	What type of data will be created? And possible data formats	Data from this WP will be mainly reports and presentations. So .docx and .pptx formats will be created.
	What types of metadata will be there?	No metadata is envisaged for this WP at this point of time.
	What past data from previous research projects will be used?	No past data from previous projects will be used.
WP2	What are the data sources and where do they come from?	The data will mainly come from literature, market studies, test facility design, experimental results, and findings/practices from BIOMETHAVERSE partners and also from other research organisations outside of the consortium.
	What type of data will be created? And possible data formats	Database with biomethane production pathways description (on MS Excel, .xlsx files).

Co-funded by the European Union

		The delivery of the 'demonstrators' blueprint' illustrating the design of the demonstrators will be defined in a later stage, both in terms of content and format, depending on the confidentiality needs and IPR. Modelling will be performed via Aspen data generated will be from process simulation outputs, to be made available in .xlsx, but also as scientific publications when possible.
	What types of metadata will be there?	The following types are envisaged: • Literature register, list of KPIs, list of key parameters. • Register of data will be available for each demonstrator.
	What past data from previous research projects will be used?	It is possible that the demonstrators will make use of results from previous projects. It is up to the pilot leaders to decide whether to disclose previous results.
WP3	What are the data sources and where do they come from?	Primary data on the technologies for the LCSA from within the consortium collected mainly from partners and integrated with literature data and modelling. Primary and secondary/background data from literature according to studies and publications on relevant topics as well as from existing LCA databases environmental: Ecoinvent, Sphera, social: e.g., SHDB, PSILCA. Background data will be obtained from literature and commercial (to be defined) databases. These background data from commercial databases cannot be disclosed, but the processes descriptions, the inventories and models built, and outcomes of the analysis, will be made public.
	What type of data will be created? And possible data formats	Two main data types will be created here: Life cycle inventory models in Excel and GaBi, data formats for life cycle inventory models are Excel xlsx and/or CSV. Life cycle impact assessment results will be mostly processed in Excel, including plots and tables.
	What types of metadata will be there?	The life cycle inventory models will include the standard meta data according to GaBI and Ecoinvent data quality guidelines. The life cycle impacts assessment results will not have any structured meta data. A detailed documentation of the inventory analysis and life cycle impact assessment will be compiled, explaining the data sources and calculation steps of the LCSA in reports/deliverables. All lifecycle inventory models according to the standard meta data format for life cycle inventory modules complemented with a comprehensive documentation in the form of reports/deliverables. The meta data will help in discovering new data and this will be done via literature research and exchange with consortium partners or partners within the LCA community.
	What past data from previous research projects will be used?	None. The data from the report 'Solid and gaseous bioenergy pathways: input values and GHG emissions:

Co-funded by the European Union

		Calculated according to methodology set in COM(2016) 767: Version 2' may be used to calculate the GHG emissions according to the RED2 sustainability criteria.
WP4	What are the data sources and where do they come from?	Data to assess replicability factors: specific (tech & economics) general (regulations) most relevant factors, requirements and specific variables affecting replicability. Ideas and inputs from project partners will be collected to define, in a participatory way, the most relevant factors, requirements and specific variables affecting the replicability of the different solution. Data collection step will be carried out through specific questionnaires that will be prepared and addressed to demonstrators and research partners to collect all necessary information. Once all variables are complete, the methodology will be applied, with the calculation of a replicability index for each demonstrator.
	What type of data will be created? And possible data formats	A diagram with every solution as a point will be created, through a mathematical approach (represented by a cartesian diagrams built on variables dependent on factors that are Solution Variables and Context Variables. On the Replication Axis: it measures the Replication Potential [0-100%], given by the intersection between the points representing the solutions and the diagonal lines. The replicability potential chart will show the ranking of solutions from the most to the least replicable.
	What types of metadata will be there?	The application of a multi-dimensional assessment method will allow the identification of the most relevant factors that may enable or limit solutions replication. The replicability analysis is Based on the analysis of 5 dimensions: Sociocultural; Institutional; Technological; Environmental and Economic – SITEE replicability tool.
	What past data from previous research projects will be used?	No past data from previous projects will be used. However RUGGEDISED, TRANSURBAN EU-CHINA, CIVITAS DESTINATIONS project applied the SITEE methodology for the replicability analysis.
WP5	What are the data sources and where do they come from?	Data from this WP will be related to dissemination, exploitation and communication. No specific scientific data will be generated from this WP. Instead, data from all other WPs will be used here for the aforementioned purposes, ensuring the appropriate and effective exploitation of the results.
	What type of data will be created? And possible data formats	Data from this WP will be created as .docs, .pptx, .pdf., jpeg. Visual identity will include logo elements, colors, templates and fonts. Templates for the following four categories: Document, Letter, PowerPoint, and Newsletter will be created. project website will be created in an attractive and user-friendly way and updated regularly. Project leaflets in electronic and printed version in 5 languages (200 each). Project

Co-funded by the European Union

		poster and roll-up English and in the other five languages. Nine e-newsletters will be prepared with six-monthly interval Social media channels, Video Pills and Media relations will ensure the promotion of the project and its main outcomes to a large audience. Scientific publications and final booklet will be produced.
	What types of metadata will be there?	No metadata is envisaged for this WP at this point of time.
	What past data from previous research projects will be used?	No past data from previous projects will be used. The results will be exploited by identifying Key Exploitable Results (KER) and Exploitation Roadmap will be drafted based on EXPLOITT methodology.

The approval of the availability of data in an open approach will need to be sent to the PC from the actual data owners via email. For this, a consent that the data can be distributed outside the Consortium must be included in the approval email to the PC. The information indicated in Table 5 should be included.

Table 5 Data ownership attributes

Data owner	Data descritpion	Data filename and version	Consent to publish/share data outside the BIOMETHAVERSE Consortium
Who is the data owner	What the data include	Filename and depository position	[YES/NO]

3.4. Generated research data

In the framework of CINEA and Horizon EU, a step toward Open Science policies is obligatory and highly recommended. All research publications coming out of the project will be available as open access and the costs for these publications will be covered from the project costs. In most cases GOLD open access will be provided and in case the budget falls short, then GREEN open access will be provided.

The project will create at least the following categories of outputs:

- public deliverables,
- scientific publications,
- open-source contributions,
- datasets from experimentation and testing.

All public deliverables will be made available on the project's website and later after the project has ended on the EU repository.

Public data arising from the project will also be shared with other similar EU projects in order increase the project's outreach and maximise the project's impact.

Broader access to scientific publications will help in the following ways:

- Avoid duplication of effort – both financial and time resources can be saved when data on similar topics is available to the public.

Co-funded by the European Union

- Encourage collaboration among researchers – There is a high likelihood that different research group, organisations and institutes are working on similar ideas and topics across the globe. Innovations and outputs generated from a group from one side of the world when made publicly available will foster global collaboration.
- Build on previous research results – Science and technology always progresses when things are built upon, and this is very crucial for further advancement.
- Speed up innovation – With the current world need for devices with higher energy efficiency and lower emissions, making data publicly available will speed up innovation.
- Involve citizens and society – All activity is ultimately done to improve the lives of people living on this planet and hence this ultimate goal must always be kept in mind.

3.4.1. FAIR principles

Data and other research outputs will be managed in line with the EU FAIR principles (Findable, Accessible, Interoperable, Re-usable) with the aim of achieving these objectives. Accordingly, The BIOMETHAVERSE DMP follows the structure of the Horizon 2021 DMP template. It reflects the status of the data that is collected, processed or generated and following what methodology and standards, whether and how this data will be shared and/or made open, and how it will be curated and preserved.

Findability: Data that will be published on Zenodo: this public data repository typically uses a DataCite or DublinCore based metadata schema. All data, including published ones and the ones that need to remain closed access will be described with Metadata readme.txt files that includes key metadata similar to DublinCore metadata fields and where required discovery metadata related to the actual research data.

Outline for keywords used:
Standard vocabulary and nuances will be used. Specific filters will allow finding data related to this project with specific keywords. The use of keywords is a mandatory field upon data deposition in the selected data repositories and will be adhered to:

- Outline for naming conventions: Clear naming conventions consisting in X mandatory parts will be created for:
    - metadata,
    - datasets for storage
- Outline for clear versioning of documents: There is a quality process in place where all documents will be provided with clear version numbers to not only find the latest document/file that must be used but also to allow users to go back to previous versions to see what exactly was changed.

To share data with the BIOMETHAVERSE Consortium, a SharePoint has been set up by ISINNOVA in Microsoft Office365 TEAMS – a platform which provides data collaboration possibilities between ISINNOVA and external parties. Files can be viewed, synced, and shared across all Microsoft compatible devices and users are able to work simultaneously on respective documents. The SharePoint enables the following:

- It facilitates data sharing for intermediate results, and final results.
- It enables users to work with quality-controlled data sets and versions approved by data owners and providers.
- It enables both data owners and data providers to choose to which WP and task they want to contribute the data and to what level.

Accessibility: Data that will be openly accessible as described above. Raw data from scientific publications will be uploaded to public repositories and other raw data from industry partners may be available upon request provided they do not fall under confidential or IP protected data.
All metadata will be made openly available under a Creative Common Public Domain Dedication as laid out in Article 17 of the GA.
The data on Zenodo is in principle available for at least 20 years and the metadata will remain available without restrictions or time limitations.

Co-funded by the European Union

- Outline for clear versioning of documents: There is a quality process in place where all documents will be provided with clear version numbers to not only find the latest document/file that must be used but also to allow users to go back to previous versions to see what exactly was changed.

Interoperability: In order to ensure maximum interoperability of data, standard or open formats will be used to facilitate data exchange both within and outside the project. Whenever software related data is concerned, commercially available or open-source software formats will be used.
The data set for scientific publications will be linked to the corresponding DOI (Digital Object Identifier) of that publication.

Reusability: Documentation readme files and Metadata will be created in addition to the typical supporting files accompanying scientific publications.
Data that will be stored in the public repository will be made available for reuse under a Creative Commons usage license like CC BY. The European Commission announced in 2019 that it intends to join other public institutions around the world to share published documents and it has adopted CC BY4.0. The CC BY is:

- Universal – applicable to all documents
- Unrestricted – generally speaking, the only condition is attribution.
- Simple – user friendly
- Cost free – no fees involved.
- Non-discriminatory – terms of CC-BY are open to all potential actors in the market.
- Transparent – text of licences is publicly available, accompanied by supporting documents, guidelines, and other material in multiple languages.

Non-confidential data will be available for maximum reuse by other interested parties and stakeholders as it is not the intention of the project or the project Consortium to withhold any data.

3.4.2. Metadata produced from the project

Metadata creation will be and is very important in the project and will be taken seriously. Creation of metadata will help other stakeholders and researchers find, identify, and discover data. The type of metadata to be created will depend on the tasks and sub-tasks carried out in the project.

For experiments, the following details will be recorded.

- The nature of the experiment.
- The devices/apparatus/equipment used, and the accuracy achieved from these.
- The experimental procedure or protocol followed.
- The operating conditions and environmental conditions under which the experiment was performed.

It is the intention that any experimental work can be recreated, and results will be reproducible once the project has concluded.

For simulations:

- Simulation tool used.
- Initial conditions or boundary conditions.
- Scientific principles applied.
- Flowchart for code or process

It is also the intention that simulation work carried out within the project must be repeatable/reproducible by other research groups worldwide solely based on the metadata provided.

For LCSA & TEA:

Co-funded by the European Union

- The procedure followed for LCSA and TEA.
- Detailed documentation of the inventory analysis and life cycle impact assessment. Data sources and calculation steps will be outlined.
- Standard meta data formats as per life cycle inventory modules.

3.4.3. Data processing & reuse of existing data

Existing data from past projects will be studied and made use of wherever possible. Past data will also be used for comparison purposes, when possible, against new improvements done in the current BIOMETHAVERSE project. Data processing from respective tasks and sub-tasks will be as per the tools and methods available with respective project partners.

3.5. Intellectual Properties Rights (IPR)

Intellectual property is key for any research organisation or industry, or other partners involved in the project. The IPR being generated out of the project either by a single partner or by means of cooperation between multiple partners will be respected and handled as per the terms and conditions stated in the CA. Therein, the necessary protocols and tools for resolving any disputes that may arise are also defined.

Data or methodologies or processes leading to IP must be kept confidential. The relevant information can either be stored on the Microsoft TEAM SharePoint or partner SharePoint with special accessibility rights to only a select few participants. In addition, data that will lead to IPR cannot be published or disseminated in any form in any public domain.

The General Assembly, which meets once in six months, will go through the generated data (up to that date) and decide on a case-by-case basis as to which data can be released and which needs to be protected for IPR.

It is highly recommended to inform the PC when any partner(s) has intentions and going ahead with IPR filing. This information is only needed so that the PC can communicate the project's success to the outside world.

Figure 4 exemplifies the BIOMETHAVERSE IPR management approach within the consortium.

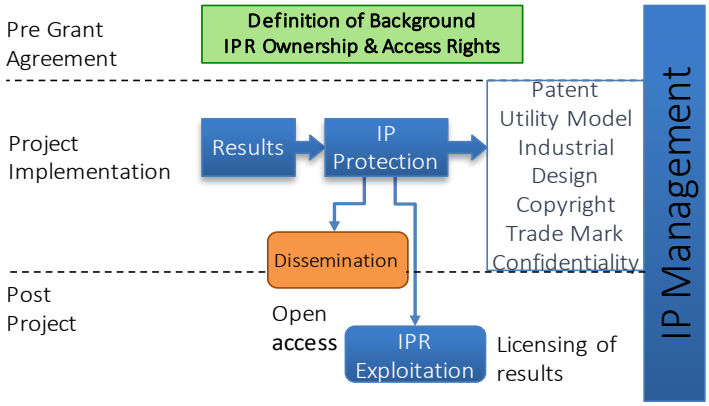

Figure 4 BIOMETHAVERSE IPR process.

4. Data storage and management

4.1. Data archiving and preservation

All data generated during the course of the project will be archived and preserved on the Microsoft365 TEAM sharepoint. This data will be available for a period of 5 years after the project has ended. Data that can be made available to the public will be transferred to respective public repositories and transferred to

Co-funded by the European Union

the corresponding EU repository as well. While sensitive and confidential data will be transferred back to the respective partners.

4.2. Which data need to be preserved?

Data generated from experimental work on short stacks, full stacks, components, sub-systems, and systems needs to be preserved. This data will be helpful for commercialisation activities, to take the technology to market readiness levels.

Data from LCSA and TEA needs to be preserved in order to provide information to relevant stakeholders and also to add to the research base for such concepts and systems. This will help policy makers in their decisions.

Data from simulations and modelling activity will also need preservation. The methodologies and logic followed for simulations will be key for further extension and improvement of activities.

4.3. What is potentially useful to others?

Data from WP2 and WP3 is expected to contribute to science and help in further advancement of the related processes, technologies, and methods. The scientific publications from the project results will help in further advancement of science in the respective areas.

Data from WP4 and WP5 will also help policy makers and industrial stakeholders in looking at the concept as a whole and assessing it from a life cycle perspective. This may help the proposed technologies to achieve higher technology readiness levels and a better market penetration.

Proving concepts and methodologies at pilot scale will provide confidence to take it to higher TRL and will pave way for bigger demonstration projects and a faster market penetration.

4.4. What has scientific, cultural or historical value?

No data is envisaged to have cultural or historic value but only scientific value. Data from WP2 and WP4 are expected to be of relevant scientific value.

4.5. What legally must be destroyed?

Any sensitive data related to any of the project partners, or third parties must be deleted and destroyed after the project period. The deletion/destruction of sensitive data pertains to removal from both Microsoft SharePoint and from local storage media of the partners.

Sensitive data that has already been used for any written material or publication (with consent from the party generating the sensitive data) need not be destroyed or deleted.

4.6. Data storage during the process

Data on the Microsoft TEAM SharePoint is regularly backed up on central and backup servers. Hence the risk of losing data is minimal. The aim of data storage during the research progress is to make data readily available to all project partners and their respective members who are involved in the project.

Critical data for each WP is to be stored on the SharePoint under the respective WP sub-area and access will be given to only those working on that particular task. The project coordinator has by default access to all data. The WP leader has the right to decide who must and must not have access to this critical data. If not, restrictions are placed then data is available for all people who have access to the SharePoint and subsequent sub-areas.

Besides the above storage location, it is highly plausible that data is also stored on the respective servers and commercial cloud storage of partners involved in the BIOMETHAVERSE Consortium.

Co-funded by the European Union

4.7. Data sharing and dissemination

Maximising the project's outreach is one of the goals within the project. Data stored on the SharePoint will be unencrypted and stored using standard character encodings to allow uninterrupted access to all project partners.

During the project phase, each party or partner is entitled to carry out dissemination activities that will promote the project or any specific topic. The dissemination activity will need consent and prior approval from both the PC and the dissemination WP leader. Each partner is highly encouraged to disseminate as much information as possible to the public unless it goes against their legitimate interests or affects other IPR arising from the project. In case legitimate interests of any partner may be harmed due to the planned dissemination, the disseminating partner needs to have written consent from the partner whose legitimate interests may be harmed, or the dissemination must not take place unless appropriate steps are put in place to safeguard the interests.

The BIOMETHAVERSE partners are encouraged to submit and publish ground-breaking results that arise out of the respective WPs and demonstrators. This will lead to both establishing a base for new scientific results and adding to further literature base of areas where results are scarce.
After the project phase, exploitation activities can be carried out for a period of 2 years, subject to the same terms and conditions as mentioned for dissemination during the project phase.
A detailed dissemination, communication and exploitation plan will be developed as part of the dedicated WP5.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

5. Data Ethics and Security

Data ethics, data protection and data security are of prime importance to the BIOMETHAVERSE Consortium and the project. The Consortium and the project will strive to ensure good research ethics and take all actions necessary to prevent any situation where personal or sensitive information could get misused.

5.1. Ethics – Personal Data

All personal information collected during the project such as names, email id's, phone numbers and others will be treated with in accordance with European data protection laws. Personal data that will be collected, will be stored, analysed, and used anonymously.

5.1.1. Dissemination, communication & exploitation activities data usage processes

It is expected that personal data will be collected mainly from WP5 'Dissemination, communication & exploitation' activities. During the course of the project, several events with third parties will be planned, such as seminars, meetings, summer/winter schools, workshops, conferences etc. It is also possible that personal data may be collected during the execution of other technical WPs.

5.1.2. BIOMETHAVERSE mailing list

The BIOMETHAVERSE list of contacts is shared via Microsoft TEAMS as xls file (WP1 files) and relies on a general mailing list for all BIOMETHAVERSE partners/contacts and one with the key contacts / WP leaders.

Additions or removals to the mailing lists are managed directly by one partner (ISINNOVA). The purpose of this list is to keep a well organised list of contacts for the BIOMETHAVERSE communications and access is restricted only to BIOMETHAVERSE Consortium partners. The mailing lists will be erased after the project end and not maintained after the project end. Any person has the right to opt out of this list by direct email to the PC (ISINNOVA).

Co-funded by the European Union

5.1.3. Meeting related material

This relates to any document created and used for the purposes of BIOMETHAVERSE meetings. These may relate to agendas, presentations, minutes, signature lists or any other internal document created for the purposes of BIOMETHAVERSE meetings. All these documents will be created and maintained for internal purposes of BIOMETHAVERSE and only F BIOMETHAVERSE partners will have access to them at the BIOMETHAVERSE repository/SharePoint under the meetings section. They will be kept for up to 5 years after the project end.

5.1.4. Workshops/Conferences and Training sessions

These data relate to the creation of workshops, agendas, programmes, participants' lists etc and in general dissemination material related to BIOMETHAVERSE organised workshops. Regarding the external publication of these, it is considered that this material can be fully anonymized so that it excludes personal information from the presenters/participants in the related programmes/agendas that will be shared publicly. For the parts of the related material that will be used for the workshop organisation internally to BIOMETHAVERSE, the related files will be stored in the BIOMETHAVERSE repository/SharePoint under the section workshops.

5.1.5. Newsletters, social media, and other dissemination material

Unless otherwise expressly specified in the GA, ISINNOVA (leader of WP5) shall be responsible for the personal data processing carried out for Project dissemination purposes. To this end, Controller shall:

- Collect and keep all relevant personal data (including lists of contact details), or copies thereof;
- Monitor relevant communications;
- Address to Project Partners instructions and guidelines on Project dissemination activities (including any EU or other state guidelines, whenever available);
- Inform Project Partners of any policy or legal requirements reviews and changes.

5.1.6. Usage of cookies (in BIOMETHAVERSE website)

If in the BIOMETHAVERSE website the usage of cookies is needed, a related pop-up window informing the user is present, prompting the user to accept (or not) the conditions under which her/his personal information are stored.

EBOS Data Protection Officer (DPO) to handle such matters and make sure subjects are informed and they consent when it comes to sharing their personal data.

5.2. Informed Consent Procedures

Data confidentiality and integrity will be we dealt at different levels.

- Data at rest – This refers to any data stored on the SharePoint or on local servers of respective partners. This data will be protected by the respective IT policies of the corresponding partners.
- Data in transit – secured by means of safe data transfer mechanisms such as TLS (Transport Layer Security)

The individuals or group of individuals whose data will be used or stored, will be informed comprehensively about the intent use of the information provided by them and will have to give prior approval in written form for use of their data for this research and scientific purpose.

5.3. Legal Aspects

Some of the legal aspects with respect to data management is discussed in this section.

Co-funded by the European Union

5.3.1. Embargo periods

Theses arising from master programs and PhD programs, from students who will work on BIOMETHAVERSE project topics, will be subject to an embargo period of 2 years after the project has finished. The embargo period and conditions only come into force if the content of the theses:

- Is confidential to one or more partners.
- Has sensitive or novel information or knowledge that is new for the project.

When needed, the partners should also adhere to the GA and to a longer embargo period be needed.

5.3.2. Sensitive data

Some of the tasks in certain WPs delve into fundamental science aspects and other novel engineering topics. It is highly probable that a lot of sensitive data will be generated from these WPs. Access to sensitive data to other partners will be with the sole consent of the party generating that data. When using sensitive data for any outreach activities, the data must be filtered and be made as general as possible in order to protect the IPR of the generating parties.

5.3.3. Data leaks

It is the duty and responsibility of the PC and every Lead from respective partners, to ensure that no data is leaked for no reason whatsoever. Examples of data leaks include, but not limited to, the following:

- Accidental publication or release of a sensitive deliverable online.
- Sensitive data from a process or partner being used without consent or published online.
- Project specific data being used or stored on personal computers and servers.
- Project specific data being used as input for another process or project without informed consent.

Data leak issues will be jointly decided by the GA where the PC will call for an emergency meeting should such an issue arise, and joint action will be taken accordingly.

5.4. Ethical aspects

The project does not involve research on humans or animals of any living beings and hence no ethical issues on this front is foreseen or expected.

5.5. Security

Data generated from the project is stored securely on the Microsoft TEAMS SharePoint. The IT principles of consortium provides for and ensures safe and secure storage of data. Data can be accessed only by project partners.

The PC maintains the SharePoint and decides which partner has what level of access to each area of the SharePoint.

Accounts with private email domains such as @gmail.com or @yahoo.com will not be granted access to the consortium share point.

Sensitive information from partners will be kept within reach to either the project Consortium or to a limited number of personnel working on that specific task or sub-task.

All project members will be responsible for ensuring that project-related data are safely and securely handled and stored during the entire course of the project, in compliance to relevant EU data protection regulations.

Co-funded by the European Union

All data on the consortium sharepoint are automatically and regularly backed up, with very low probability for any data loss.

Co-funded by the European Union