



D5.3

Data Management Plan, Open Science and protection strategy

Consortium:



Bay Zoltán
Nonprofit Ltd.
for Applied Research



EUROPEAN ASSOCIATION
OF APPLIED UNIVERSITIES



MOHOLY-NAGY
művészeti egyetem
university of art and
design budapest



business
upper
austria



Metropolia
University of Applied Sciences



Funded by
the European Union

 www.engage4bio.eu

 info@engage4bio.eu

@Engage4BIO



D5.3

Data Management Plan, Open Science and

DELIVERABLE TYPE

Report

MONTH AND DATE OF DELIVERABLE

Month 6, 30/03/2023

WORK PACKAGE

WP5

LEADER

ZSI

DISSEMINATION LEVEL

C-UE/EU-C

AUTHORS

Hofer M., Feichtinger J.,

Programme

HORIZON
EUROPE

Grant Agreement

101059565

Start

Oct.2022

Duration

36
Months



Contributors

NAME	ORGANISATION
Margit Hofer	ZSI
Judith Feichtinger	ZSI
Maria Schrammel	ZSI

Peer Reviews

NAME	ORGANISATION
Serena Fabbrini	APRE

Revision History

VERSION	ORGANISATION	REVIEWER	MODIFICATIONS
1	02/03/2023	Feichtinger	Corrections
2	7/03/2023	Schrammel	Corrections
3	22/03/2023	Fabbrini	Review and editing, V3 sent to the Coordinator

The information and views set out in this report are those of the author (s) and do not necessarily reflect the official opinion of the European Union, neither the European Union Institutions and Bodies nor any person acting on their behalf.



Index of Contents

1	Introduction.....	8
2	Data Summary.....	9
2.1	Purpose of data collection	9
2.2	Types and format of data.....	9
2.3	Reuse of existing data and origin of data.....	10
2.4	Format and Size of the data	10
2.5	Data utility.....	11
3	FAIR Data.....	12
3.1	Making Data Findable Including Provisions of Metadata	12
3.1.1	Versioning.....	12
3.1.2	Naming.....	12
3.1.3	Metadata and keywords.....	13
3.1.4	Version numbers	15
3.1.5	Metadata	15
3.2	Making Data openly accessible	15
3.2.1	General rules regarding accessibility.....	15
3.2.2	Research data access	15
3.2.3	Sensitive Social Science Data.....	16
3.3	Making Data Interoperable	17
3.4	Increase data re-use ((through clarifying licenses)	17
3.4.1	Timeframe of the re-usability of Data	18
3.4.2	Data quality assurance processes.....	18
4	Other research outputs.....	19
5	Allocation of Resources	19
5.1	Costs of making data FAIR	19
6	Data security	20
7	Ethical aspects	21
7.1	Ethical or legal issues.....	21
7.2	Gender and diversity	21
7.3	Informed consent for data sharing and long term preservation	21
8	Other issues.....	23
8.1	Other national/funder/ sectorial/departmental procedures.....	23
9	DMP Maintenance.....	23

10	Conclusions, summary and outlook.....	24
11	Table of Tasks incl. Type, Format and Re-use.....	25
12	Appendix.....	28
12.1	Appendix A: Engage4BIO open research policy.....	28
12.1.1	Transparency	28
12.1.2	Participation	28
12.1.3	Reusability of knowledge.....	28
12.2	Appendix B: Detailed Information Sheet.....	30
12.3	Appendix C: Engage4BIO consent file for Co-Creation workshops.....	33
13	References.....	34

Index of Tables

Table 1	Data description	14
Table 2	List of tasks incl. Type, format and Re-use	25

Index of Figures

Figure 1	Open Access work flow of Engage4BIO	19
----------	---	----

Partners short names

ZSI	Zentrum Fur Soziale Innovation Gmbh
WR	Stichting Wageningen Research
APRE	Agenzia per la Promozione della Ricerca Europea
BZN	Bay Zoltan Alkalmazott Kutatasi Kozhasznu Nonprofit Kft.
EAEA	European Association For The Education Of Adults
MOME	Moholy-Nagy Muveszeti Egyetem
artEZ	Stichting Artez
CLIC	Clic Innovation Oy
TMG	Business Upper Austria – Oo Irtschaftsagentur Gmbh
MET	Metropolia Ammattikorkeakoulu Oy
UNIPA	Universita Degli Studi Di Palermo

Abbrevations

AUSSDA	Austrian Social Science Data Archive
CA	Consortium Agreement
CSA	Coordination and Support Action
DCMI	Dublin Core Metadata Initiative
DI	Digital Innovation
DMP	Data Management Plan
DoA	Description of the Action
DOI	Digital Object Identifiers
DPO	Data Protection Officer
DSI	Digital Social Innovation
EU	European union
FAIR	Findability, Accessibility, Interoperability, and Reuse of digital assets
GDPR	General Data Protection Regulation
HE	Horizon Europe
IP	Intellectual Property
OA	Open Access
OAI	Open Archives Initiative
ORDP	Open Research Data Pilot
RRI	Responsible Research and Innovation
SME	Small and Medium-sized Enterprises
URI	Uniform Resource Identifier
WP	Work Package

Executive Summary

Allowing research data to get reused by other researchers and all kind of stakeholders is a core principle of Engage4BIO. The goal of this document is to ensure that third parties can freely access, mine, exploit, reproduce and disseminate data and publications of Engage4BIO. Following the well-known strategy "as open as possible, as closed as necessary", it is acknowledged that not all data can be open.

This deliverable presents the strategy and templates for the management of data within the Engage4BIO project, its Data Management Plan (DMP), for the Coordinator and beneficiaries. During the lifecycle of a project the generated data may evolve in terms of volume, type, speed, sources, analysis processes and so on. It is therefore important to consider a DMP as a dynamic framework that might change during the project.

1 Introduction

Engage4BIO is a Coordination and Support Action (CSA) that aims to strengthen circular, sustainable bioeconomy and regional development through engaging quadruple helix actors - taking into account their diversity of societal, economic and cultural perspectives, in five regional bio-based systems (hubs) in processes of design thinking, co-creation, (re)training and skills development. The project focuses on launching actions for the deployment of local bioeconomies and new ways to govern the required societal transformation and, at the same time, engage citizens through awareness raising and education on sustainable production, consumption and lifestyles. Based on their regional specificities (availability of biomass feedstock, level of technology, planetary boundaries and social capital) activities will be co-created, performed and (re-) shaped to their best advantage ensuring the inclusiveness of all actors.

Given the high involvement of the different stakeholders in the planned workshops and actions, Responsible data management is crucial for this project. The Data Management Plan (DMP) provides structured guidelines for this process and will ensure a responsible handling of any data collected. The DMP covers the lifecycle of data, including the process of collection to storage, analysis and preservation. In this DMP we provide a basic description of what kind of data will be produced and collected during the course of Engage4BIO, and details about what will happen to the data both during this project and after it has been completed.

Engage4BIO is part of a pilot action on open access to research data and is thus committed to providing access not only to project results and processes, but also to data collected during that process.

While the general policy of the Engage4BIO project is to apply “open by default” to its data, privacy issues are treated with specific care. Legal rules on anonymity are highly relevant and need to be agreed upon with each of the participants. Most of the data in Engage4BIO will be anonymised data, however in case of a doubt, the data privacy of our participants always prevails over open data policy. When dealing with personal data Engage4BIO is making sure that all GDPR rules are followed.

Overall, the leading framework for our data management is the GDPR in terms of personal data protection, FAIR data for the handling of non-personal data and Open Access following the European Union’s Open Science Guidelines.

2 Data Summary

A Data Management Plan (DMP) is a key document of projects that collects or handles original data. This DMP describes the data management life cycle for the data to be collected, processed and/or generated by the Engage4BIO project, following the Guidelines on FAIR Data Management in Horizon 2020 provided by the European Commission¹. As part of making research data findable, accessible, interoperable and reusable (FAIR), a DMP should include information on:

- what **type of data** will be collected, processed and/or generated
- the **handling of original data** during and after the end of the project
- which **methodology and standards** will be applied
- whether data will be shared/made **open access**
- how data will be **curated and preserved** (including after the end of the project)

The DMP is considered a living document and will be updated if needed due to changes in the consortium policies, research methodology or other significant development which affect the management of data. Therefore, we will implement this deliverable as a document that can be updated by the partner organisations at any time in the course of the project. In preparation for the mid-term review we will also reflect on the status of the DMP with the whole consortium and address any possible needs for updating the document and the processes described in it.

Engage4BIO is influenced by the integrative concept of Responsible Research and Innovation (RRI), which clearly goes beyond a pure ethical aspect. The importance of the RRI approach for the project is to align the co-creation process, which is an integral part of most activities, in a reflective way that is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other.

2.1 Purpose of data collection

Part of the data description will always be the aim of the data produced and how it is linked to the objectives of the project. Similarly, a short description of the process of the data collection will be applied so as to help to characterize the specific data.

In Table 1 page 13f. of this document, shows the Engage4BIO data in as much detail as we can at this stage of the project. We remain cognizant of all the points outlined above, and we will update Table 1 in case of additional data being handled in any of our research activities.

2.2 Types and format of data

Due to the interdisciplinary nature of Engage4BIO, a combination of several research science methods will be used to collect the research data. Data types will vary between the different Work Packages (WPs) within the project. This is accurate

¹ https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

for the data sets collected, generated and even disseminated. Therefore, each WP will be handled separately to ensure accuracy as follows:

Observational: data captured through observation of an activity and/or a behaviour in real-time (and, therefore, usually irreplaceable). The collection will require human observation, ethnography, and potentially also instruments, including surveys, interviews, focus groups, to record the information.

- Experimental: data obtained in controlled situations through the active intervention of researchers, who design an activity (called the experiment) to collect data. Such data can be reproduced if needed, but it may be time-consuming and expensive.
- Simulation: data generated through reproduction of a real-world system, usually used to determine what would happen in certain conditions.
- Derived or Compiled: data derived upon transformation of existing datasets to create new data. Such data can be reproduced if lost, but it is very time-consuming and it could be expensive.
- Reference or Canonical: to data used to categorize and/or classify other data. They could be static or changing over time.
- Event-related: data collected during events. As they are derived from real-time events, the data are irreplaceable. If they include personal data of participants, just as in the other types of data, the GDPR will be applied.

While most of the research data can be assigned to one of these categories in some cases there could also be an overlap. Some data can belong to more than one category depending on the specificities of the data itself (e.g. event-related data could also be considered observational data under certain conditions).

2.3 Reuse of existing data and origin of data

For the interpretation of the research results, statistics and survey results, any existing data, may or will be used. Engage4BIO will not commission or purchase any survey data from third parties.

Mostly it is expected to rely on using available public data sources as found in public repositories and online. When applicable, some of the WPs will be re-using data already available from project partners or externals when the data is proved to support the ongoing research or demonstrations operation. All the data re-used for the fulfilment of the Engage4BIO project will be integrated in the data set that is going to be made publicly available. Until now, Engage4BIO is not planning to commission or purchase any data in forms of surveys from third parties.

Engage4BIO will produce its own data for the purpose of the research proposed to the Commission. The use of other data for this purpose is not planned and therefore not part of the research data collection plan.

2.4 Format and Size of the data

Our data can have various formats: text, numeric, multimedia, models, software languages, codes and design, discipline and instrument specific, etc. It is important to describe the format of the data in order to facilitate its accessibility and

interoperability. Preference will be given to open and standard formats. Likewise, data management experts advise to give indications about the size of the data collected in a DMP. Similar to the format of the data, size is relevant for future data-sharing and making data accessible. In the course of the time information about the size of the managed data in Engage4BIO should be revised and updated accordingly.

The exact volume of the data, which will be deposited in a respective folder of the cloud2 server - the repository for research data and publication at ZSI - is not known at this stage. Based on the types and formats of the expected project research data and publication (deliverable), the estimated volume is between 100 –200 Gigabytes (GB).

2.5 Data utility

Engage4BIO is a multidisciplinary innovation project funded under Horizon Europe, where several stakeholders from different disciplines joining to elaborate into a collaborative co-development; therefore, the collection of research data and materials will be useful for stakeholders in bioeconomy:

- Academic researchers
- Practitioners
- SMEs and industries
- Communities
- Regional development institutions
- Innovation accelerators
- Bioeconomy Policy
- General public

3 FAIR Data

FAIR data means that the data are findable, accessible, interoperable, and reusable. In this section we elaborate on the measures the project is taking to generate FAIR data during the course of the project and to ensure FAIR use of the data beyond the project timeframe. This makes sure that the data is well formatted in order to exploit the ability of machines to automatically find and use the data, in addition to supporting its reuse by stakeholders (Wilkinson et al. 2016). According to FAIR principles, data should be:

- identified in a persistent manner;
- properly described with enough metadata;
- stored so that both humans and machines can easily access it;
- well structured so that it might be integrated with other data sets;
- clearly presents well defined licence or terms of use.

3.1 Making Data Findable Including Provisions of Metadata

3.1.1 Versioning

The datasets produced in Engage4BIO will follow good research practice by following a set of naming and versioning conventions, which are very useful for collaborative projects. A version control system will be applied to Engage4BIO sure that the data is organized and we can track back among different versions of the datasets. Such a version control system, which is also sometimes referred to as a revision control system, tracks incremental versions (or revisions) of files over time. Cloud2, which is the main data-sharing system running on a ZSI server, supports a simple, automated version control system for files. Through use of this version control system, Engage4BIO will minimize the risk of losing information after modifications in collaborative processes where many Engage4BIO consortium members are working with the data. All consortium partners have access to the ZSI Cloud2.

In addition to the automated versioning via Cloud2, we are also applying version numbers in the file names themselves (see naming convention below).

3.1.2 Naming

Engage4BIO follows an established good practice when it comes to naming of deposits. This includes proper file naming which is sufficiently descriptive so that the content can be assumed irrespective of where the file is stored. A minimum possible number of characters must be kept when naming the files.

For the naming of our shareable data we will use file names that are informative and useful for both humans and machines in order to Engage4BIO the data accessible and findable. Names must be meaningful, on the one hand, in order to Engage4BIO it easy for others to understand what the file contains and how it should be used. On the other hand, names should use deliberate delimits in order to be machine readable.

A common approach is using characters like “_” and “-” to delimit units of metadata in the file names. For this project, we will aim to use “-” to separate words we want to glob together, and “_” to separate different information within a file name. Blank spaces, punctuation, upper-casing of entire words, or special characters (e.g. characters such as \$, @, %, #, &*, (,), !, etc.) which may have meanings in programming languages, will be avoided.

As mentioned above, the file-naming should also always include a version number to make sure others who want to make use of any dataset know that they are dealing with which version. An example of an appropriate name for a data file is: *e4b_wp1_Collection_gaps_v1.xml*

3.1.3 Metadata and keywords

Descriptive and substantive (i.e. how the data should be read or interpreted) metadata will be elaborated and described in a readme.txt file complementing each dataset (see Table 1). This detailed documentation should help interested users to clearly understand and reuse the data.

Title	Description
Creator(s)	Main person(s) involved in producing the data.
Title	Name or title by which the dataset is known.
Contributor(s)	Person(s) or organization(s) responsible for making contributions to the dataset.
Publisher	Holder of the data (e.g. in archives), or person or institution who submitted the work.
Data ownership	Indication of primary data controller (person) and indication of additional data ownership, especially in the case of shared ownership.
Year of publication	The year when the data was or will be made publicly available.
Data created	The date when the resource itself was put together (could be a date range or a single date)

<p>Description</p>	<p>Concise description of the contents of the dataset. Describe the research objective, type of research, method of data collection, type of data and storage location of data (especially in the case of data not available on Zenodo).</p>
<p>Subject</p>	<p>Subject, or key phrase describing the resource.</p>
<p>Temporal coverage</p>	<p>The dates to which the data refer (the year or beginning and ending dates).</p>
<p>Spatial coverage</p>	<p>The geographic area to which the data refer (e.g. municipality, town/city, region, country). The geographic coordinates of the area may be included, if desired.</p>
<p>Identifier(s)</p>	<p>Zenodo automatically assigns a DOI to a dataset once the entire deposit procedure has been completed. In some cases, a dataset may be known by one or more (persistent) identifier(s).</p>
<p>Language</p>	<p>The primary language of the resource.</p>
<p>Link to related publication(s)</p>	<p>Include the web addresses or DOIs for any publication, important internal reports or other datasets that are related to your dataset.</p>
<p>Keywords</p>	<p>Keywords that associate with the data resource.</p>

Table 1 Data description

3.1.4 Version numbers

As Table 1 shows, each data resource will have a unique identifier, e.g. a Digital Object Identifier (DOI). As we are committed to sharing all our open data on Zenodo, a DOI will be generated automatically and this will serve as our main data identifier. For some data, such as the open educational resources of WP3, we will apply a more detailed metadata schema, based on the DataCite Metadata Schema².

3.1.5 Metadata

The stored data are described by using the standard metadata schema Qualified Dublin Core³. To meet the requirement of the Deposit service partners, e.g. the DOI registration agency DataCite⁴ or Open Access initiatives like OpenAIRE⁵, there are some additional qualifiers.

3.2 Making Data openly accessible

3.2.1 General rules regarding accessibility

Engage4BIO consortium members support the objective of opening up access to research results. Where applicable, the Engage4BIO members are going to use open source licenses for software and non-exclusive licenses for other intellectual property rights such as copyrights, patents or standardization proposals. Also, access to the collected data will be open and free while adhering to privacy and personal data protection. Peer-reviewed scientific publications that result from the project will be provided via open access (“gold” model). We will also consider submitting our publications to Open Research Europe⁶ and Open Access European Commission⁷, the open access publishing platform for scientific articles launched by the European Commission as a free service for HE beneficiaries.

Issues related to protecting intellectual property are highly important when citizens and volunteers are involved in the work. The Consortium Agreement (CA) (signed by all consortium members before the project start) is the place where the details of management of intellectual property rights, including open licenses, between partners are defined. The CA also includes a list of pre-existing know-how (according to the description of work) and knowledge as well as the major principles on exploitation and dissemination issues. The consortium members have already agreed that access rights on the pre-existing know-how needed for carrying out the project shall be granted on a royalty-free basis.

3.2.2 Research data access

The main part of raw and processed data generated by the project will be accessible to all consortium partners, i.e., all data except where there is sensitive data that requires being kept only by the partner having collected the data. The main

² https://schema.datacite.org/meta/kernel-4.4/doc/DataCite-MetadataKernel_v4.4.pdf

³ <https://www.dublincore.org/> (last accessed on the 03.03.2023).

⁴ <https://datacite.org/> (last accessed on the 03.03.2023).

⁵ <https://www.openaire.eu/> (last accessed on the 03.03.2023).

⁶ <https://open-research-europe.ec.europa.eu/>

⁷ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

workflow for data sharing starts at the work package (WP) level, where each team is responsible for respecting ethical procedures at all times during the data gathering and processing steps. The WP/task team members are also responsible for any data anonymization or pseudonymization, if applicable. The data accessible to the whole consortium will be stored on the shared Cloud2 server provided by ZSI. Only members of the consortium, after validation by the coordinator team (Judith Feichtinger, Margit Hofer and Maria Schrammel), can access this space. The public data will be accessible via Zenodo.

Confidential data and data collected for internal purposes will be stored in the secure facilities of the organization responsible for collecting the data and will be retained for two years after the end of project. Data containing personal identifiable information of individuals will only be shared, when necessary, based on a confidentiality agreement (see consortium agreement). A data exchange form, signed by the partners exchanging sensitive data, documents the data exchange between these partners. A specific form will be provided by the coordinator. Details concerning the ownership, transfer and dissemination of project results are defined in section 8 of the Consortium Agreement and shall be followed accordingly. The relevant rules of the Grant Agreement, in specific Article 26(2), Article 29(1) and Article 30, are also relevant and apply accordingly.

3.2.3 Sensitive Social Science Data

In Engage4BIO we are dealing with quantitative and qualitative social science data which is often context-sensitive data that might include personal data. GDPR applies for any such personal data, defined as follows:

'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person

(Art. 4(1), GDPR).

Thus we have to pay careful attention not only to avoiding direct identifiers in our data, but also indirect identifiers that could make persons identifiable through aggregation of information. For special categories of personal data it may even be necessary to provide for certain suitable and specific measures to safeguard the fundamental rights and the interests of the data subject. Protection of data subjects according to GDPR rules will take clear precedence over our commitment and intention to contribute as much as possible to open research. If necessary, we may also apply data encryption to make sure sensitive personal data is protected.

There are certain measures that can be taken to allow even highly qualitative social scientific data to be made available and shareable, and we will make use of these measures as far as possible. This includes e.g. clear procedures for informed consent, anonymisation, pseudonymisation and/or access restrictions.

We will eliminate direct and indirect identifiers through one or more of the following:

- Deleting certain information
- Substituting personal identifiers with pseudonyms (more relevant for qualitative data, e.g. interview transcripts)
- Aggregating the information (more relevant for quantitative data, e.g. numeric dataset)

The Austrian Social Science Data Archive (AUSSDA) offers support for data pseudonymisation via pseudonymisation checks. As the ZSI is a collaboration partner of AUSSDA we can make use of this service for our Engage4BIO research data. Still, if we anonymise the data to make it shareable we have to ask ourselves the question of how far the sharing of this qualitative and often highly context sensitive data makes sense and how far this data could then still be reused by others.

If we allow restricted access to the research data to enable data sharing for scientific purposes, we will follow the principle of “as open as possible - as closed as necessary”. Such restricted access could be granted to a predefined group of users or to certain accounts and under certain conditions. Within the consortium we also have a data sharing agreement in order to exchange research data within the consortium. Such restricted access would only be granted to consortium-external researchers with the consent of the research subjects. In cases where datasets can be made freely available as they do not include any personalised data we will make them available with the corresponding metadata on Zenodo⁸.

3.3 Making Data Interoperable

The interoperability of the data mostly depends on the metadata that is provided with the data. Table 1 above described the general metadata that we will provide with each of our datasets. In addition, Zenodo requires metadata for the description of any data shared on the portal. This metadata uses a vocabulary that follows the FAIR data principles and that offers export options to international standard formats, such as Dublin Core or the Datacite Metadata Schema, to Engage4BIO the data interoperable. To allow inter-disciplinary interoperability we plan to use standard vocabularies for all data types present in your datasets. However, there may be uncommon vocabularies being used that stem from the Engage4BIO-specific context in which we are operating. These context-specific vocabularies may require provision of some additional explanation with the data, in order to Engage4BIO them interoperable across disciplines.

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

The main users of data produced in the project will be the consortium partners. Following an open and participatory approach towards data sharing and re-use, we also consider participants in our project activities as potential users of the collected data. All data will be classified as open, sensitive or closed, with the access and re-usability depending on the classification. For re-use of the data the following applies:

⁸ <https://zenodo.org/>

- Open data will be freely available, subject to the user acknowledging it through citing the name and creator of the dataset. Permission from the main data creator (main researcher) or contributors is not required.
- Sensitive (or confidential/restricted data) may be made available by the data creator after any identifying information has been removed. Reuse of this anonymised/pseudonymised data may be possible, if cited and if agreed to by the dataset creator and its contributors.
- Closed data are not available for sharing or reuse.

The open data generated in this project will carry the Creative Commons Attribution (CC-BY) license so as to permit its reuse. As previously mentioned, all relevant open data will be preserved and archived in Zenodo. As far as we can say at this point, all data and items on Zenodo will be retained for the lifetime of the repository. There is currently no intention to remove the shared data from the repository at any given time. Zenodo is hosted by CERN which has an experimental programme defined for the next 20+ years, offering long-term accessibility of the data.

3.4.1 Timeframe of the re-usability of Data

All the research data stored in the different repositories are intended to remain usable for an indefinite time. For example, Zenodo explains the data retention period to be 'items will be retained for the lifetime of the repository' this is currently the lifetime of the host laboratory CERN defined by at least 20 years.

In case of the linked graph database instance with wikibase, it is supported for a period of three years after the end of the project.

3.4.2 Data quality assurance processes

Currently internal review processes of publications are described. In addition to that, data quality assurance processes include:

- Scientific papers: peer review, well described.
- Own publication: internal review process and adherence to professional practices of scientific and communication writing, ad hoc processes relying on the excellence of the partners.

In addition, data quality assurance processes and measures are described well throughout this DMP concerning:

- Standard scientific procedure and guidelines for data collection
- Conformance with research ethics
- Data security
- Metadata
- Data documentation and organization
- File naming versioning, etc.
- Repository selection, deposition and preservation

These measures to ensure data quality will be applied in all the different phases from data collection, data deposition to long-term data preservation, access and reuse.

4 Other research outputs

At present, we are not aware of other research outputs that could impede individual Engage4BIO project partners not providing FAIR data as described in this DMP.

5 Allocation of Resources

Engage4BIO is part of the Open Research Data Pilot (ORDP), a pilot action on open access to research data, which requires projects to define and execute a Data Management Plan (i.e., the plan outlined in this document). Engage4BIO's open data management is part of Task T5.4 of the Management WP, under the lead of the coordinator ZSI. All other consortium partners are committed to contributing to the data management in their roles as WP leaders. The workflow of how data will be published is depicted in Figure 1 below.

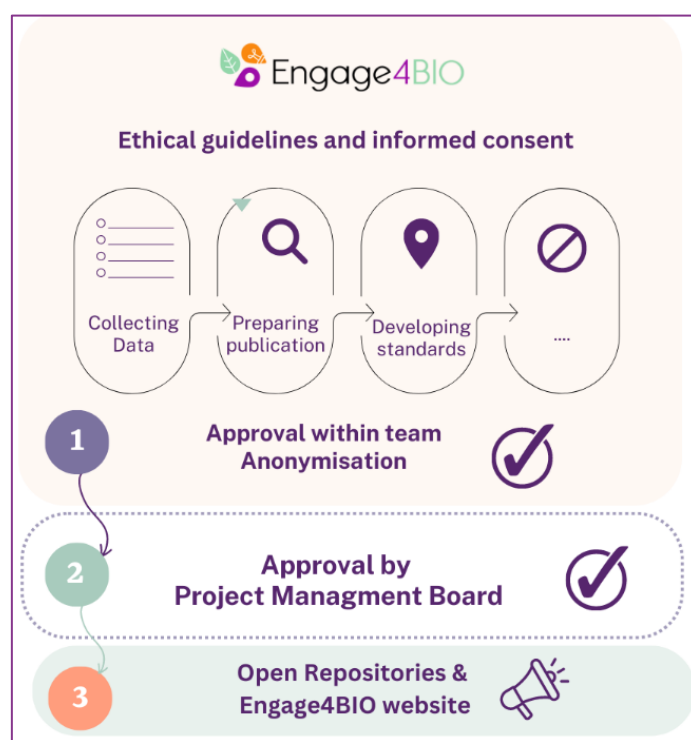


Figure 1 Open Access work flow of Engage4BIO

This Data Management Plan may have to be revised during the course of project activities. As the co-design approach that Engage4BIO is following is a rather dynamic methodology, it is not possible to clearly specify all data sources and collected outcomes from the beginning.

5.1 Costs of making data FAIR

As an EU Horizon Europe Support Action, the activities in Engage4BIO are one part of the overall activities taking place within the project. The costs for making any of our research data FAIR can be covered by the project budget during the official funding period. According to the Grant Agreement, costs related to open access to research data are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions).

6 Data security

All research data from this project will be stored on the collaboration space Cloud2, which is running on a ZSI server and is fully GDPR compliant. ZSI is running its own IT infrastructure, and its IT staff takes great care to ensure data security and protection via these measures, among others:

- in-house servers controlled exclusively by ZSI IT staff
- services running in a demilitarized network zone behind a redundant firewall
- resource isolation for services through hardware nodes and/or virtual machines
- regular updates of system and application software
- timely installation of security patches from OSS suppliers
- daily backups, off-site backup on encrypted hard disks, monitoring and logging
- web server hardening
- a password policy
- consultation and awareness-raising with ZSI staff on data protection issues
- precautions for emergency scenarios
- compliance with GDPR

The Cloud2 collaboration space has a versioning service so that data can also be restored in case of any accidental deletion. Consortium partners may also store the data they generate on their own secure servers, making sure that secure storage is guaranteed. The conditions for transfer of any sensitive data are described above. Under no circumstance will Engage4BIO collect, store or process private information, etc. without the explicit and written permission of the respective persons. Engage4BIO will comply fully with data protection provisions as set out in the currently applicable Data Protection Directive EC 95/46, the General Data Protection Regulation EU/2016/679 (GDPR) and the e-Privacy Directive EC 2002/58 (confidentiality of information, treatment of traffic data, spam and cookies).

7 Ethical aspects

7.1 Ethical or legal issues

Research must respect ethical standards and fundamental rights to respond to societal challenges. During the lifetime of the project, Engage4BIO will continuously follow an approach that ensures an appropriate level of ethical sensitivity, meaning that any changes and implementations will be viewed from an ethical perspective and (where needed) actions taken accordingly. Those consortium members that are research performing organizations have internal ethical boards and will seek their approval for any data sensitive activities.

7.2 Gender and diversity

Gender equality and gender diversity The Engage4BIO project fully supports the European Union policy on equal opportunities between women, men and gender-diverse persons. To this end, participation of all genders will be continually encouraged and supported. Project staff of all consortium members are composed according to the principle of equal opportunity and all member organizations strongly enforce an equal opportunity policy in their human resources selection processes. In addition, we are aware of inherent gender biases that can be present Engage4BIOspaces and will follow a gender-sensitive methodology, following gender-sensitive principles. Any dataset openly shared will as far as possible be controlled for bias, as well as for research ethics and integrity.

Enhancement of current education processes to better equip future Engage4BIO, entrepreneurs and society as a whole with the necessary competences to participate in collective and responsible processes. Of specific importance for Engage4BIO is the aspect of learning and knowledge sharing, within the specific Engage4BIO spaces, among and across them, and beyond the Engage4BIO community. It is critical that we transfer our findings into useful knowledge resources for the target communities (e.g. Engage4BIOs and innovators). Thus, specific activities in WP2 are dedicated to co-designing and testing measures to support competence building, especially skills that are highly relevant for the future such as critical thinking, sustainable making and digital fabrication skills.

7.3 Informed consent for data sharing and long term preservation

The standardize informed consent forms used by all Engage4BIO project partners includes consent for data sharing and long terms as well as:

- Project introduction
- Purpose
- Eligibility
- Type of intervention (interview, survey, questionnaire, co-creation session, etc.)
- Statement for voluntary participation
- Procedure
- Benefits
- Reimbursements

- Confidentiality
- Sharing of research finding
- Collection and preservation of personal data
- Rights to refuse or withdraw
- Contact persons

All content which is stored in Cloud2 is anonymized according to the policy for safeguarding good scientific practice of ZSI GmbH.

Anonymization of personal data: When submitting a data file in Cloud2 the submitter confirms that the research data which he/she has submitted do not contain any personal data. If personal data are contained, they are anonymized completely according to canonical standards and the human subjects have consented to the data collection as well as to the publication of the (anonymized) data.

Personal rights: When submitting a data file in Cloud2 the submitter also confirms that by submitting personal metadata (name and surname of the participating scientists) he/she acts in consent with all persons whose data he/she enters. He/she also confirms that he/she doesn't violate any personal rights by omitting the name of any participating scientist.

8 Other issues

8.1 Other national/funder/ sectorial/departmental procedures

At present, we are not aware of issues that could impede individual Engage4BIO project partners not providing FAIR data as described in this DMP.

The partners will compare the DMP to their institutional policies and guidelines when it comes to research and data management, data protection and security. This includes discussions with the responsible personnel within their organizations. If any issue is perceived, partners are encouraged to report it to the project coordinator by providing adequate information to be able to understand the situation/case and then suggest proper solutions. These perceived issues together with the suggested solution and decisions taken will then be covered in the upcoming updates of Engage4BIO DMP.

9 DMP Maintenance

The Engage4BIO Data Management Plan is being maintained and updated by the project coordination. This will take place in consultation with the project research institution.

Engage4BIO DMP is considered a living document and will be updated whenever needed. Reason for DMP updates include any significant project development that effects directly the management of data, changes in consortium policies, research size and methodologies.

Contact for the DMP

Dr. Judith Feichtinger, Mag. Maria Schrammel, Dr. Margit Hofer (Project Management team)

Email: e4B@zsi.at

10 Conclusions, summary and outlook

This document describes the main aspects of how this project will practice open research and how we will handle our research data. While Engage4BIO is fully committed to open science and openness as defaults, the privacy and data security of personal data coming from our research participants is of the highest priority. Templates and procedures for data collection, including for ethical compliance, are in place and will be reflected in the consortium.

As research data management is a continuous process and we can currently not foresee all eventualities that may arise through following an open co-design methodology, we will reflect on and update this DMP in the course of the project.



11 Table of Tasks incl. Type, Format and Re-use

Table 2 List of tasks incl. Type, format and Re-use

Task	Short data description & aim of the data	Type of data* (observational data, experimental data, ...)	Format of data+ (DOCX, ODF, ODT, PDF, CVS, JPEG, ...)	Reuse of any existing data
WP1	Setting the scene - concepts and content			
T1.3	Hub exchange	Data Collection in the context of WP1_Task 1.3: - Collection of Data relating to bioeconomy: Hub-to-Hub exchange	DOC/DOCX/ODT	No reuse of pre-existing data Use of data compiled by the project
ST1.3.1	Cross-fertilisation workshop 1	responses from cross-fertilization workshops; qualitative data	DOC/DOCX/ODT GIF/JPGAudio files (mp3/mp4)	No reuse of pre-existing data Use of data compiled by the project
ST1.3.2	Cross-fertilisation workshop 2	responses from cross-fertilization workshops; qualitative data	DOC/DOCX/ODT GIF/JPGAudio files (mp3/mp4)	No reuse of pre-existing data Use of data compiled by the project
ST1.3.3	Cross-fertilisation workshop 3	responses from cross-fertilization workshops; qualitative data	DOC/DOCX/ODT GIF/JPGAudio files (mp3/mp4)	No reuse of pre-existing data Use of data compiled by the project
ST1.3.4	Cross-fertilisation workshop 4	responses from cross-fertilization workshops; qualitative data	DOC/DOCX/ODT GIF/JPGAudio files (mp3/mp4)	No reuse of pre-existing data Use of data compiled by the project

WP2	Co-creation and co-design			
T2.1	Map and gap analysis in each hub	<p>Interview responses Data collection via application processes especially for sourcing stakeholders and matchmaking purposes (applicants enter information themselves):</p> <ul style="list-style-type: none"> - Contact details - Collection of SME related data - Collection of founder related data such as name, experience, expertise etc. <p>Collection of existing content and co-creation/learning and outreach materials, and content created for this project from consortium partners.</p>	<p>XLS, PDF, Audio files (mp3/mp4)</p> <p>Short Videos, print, translated format, tutorials, PDF</p>	<p>Reuse of survey data collected by WP 1</p> <p>To some extent, there will be re-use of existing contents and learning materials provided by consortium partners.</p>
T2.2	Co-creation of local bioeconomy vision and strategy approach	workshop format with respective stakeholders; informed consent needed	DOC/DOCX/ODT; XLS, XLSX	Use of T2.1
T2.3	Co-creation of guidelines for training and mentoring for adults	workshop format with respective stakeholders; informed consent needed	DOC/DOCX/ODT	<p>No reuse of pre-existing data</p> <p>Use of data compiled by the project</p>
T2.4	Co-creation workshops on knowledge gain campaigns in hubs	workshop format with respective stakeholders; informed consent needed	DOC/DOCX/ODT	<p>No reuse of pre-existing data</p> <p>Use of data compiled by the project</p>
T2.5	Co-creation workshops on governance models	workshop format with respective stakeholders; informed consent needed	DOC/DOCX/ODT	<p>No reuse of pre-existing data</p> <p>Use of data compiled by the project</p>

WP3		Implementation and Translation		
T3.1	Implementation of activities supporting regional development paths	collection of Data from regions on bioeconomy operations and their involvement in bioeconomy quantitative and qualitative data	DOC/DOCX/ODT; XLS, XLSX	No reuse of pre-existing data Use of data compiled by the project
T3.2	Implementation of training and mentoring programmes, etc.	Organise events and identify and set up a group of trainers where they will be trained on bioeconomy	JPEG, GIF; PDF	Use of T2.1
T3.3	Implementation of art events and communication campaigns	Organise art events campaigns for bioeconomy	JPEG, GIF; PDF	Use of T2.1
T3.4	Awareness raising campaign	Organise awareness and outreach campaigns for bioeconomy	JPEG, GIF; PDF	Use of T2.1
ST3.4.1	This is bioeconomy - international design awards	Data collection incl. 'names, contact details, innovative bioeconomy product	DOC/DOCX/ODT XLS, XLSX	N/A
WP4		Communication, dissemination and exploitation		
T4.2	Dissemination and communication tools and activities: Newsletter; Social media statistics Google analytics for the website	Name and contact details of subscribers	Collected in Mailchimp XLSX/spreadsheets, DOC/DOCX/ODT, PDF, JPEG, Audio and Video files [mp3 / mp4]	N/A
T4.3	Liaison with other projects, initiatives, regions and networks	Minutes and input from discussion rounds and exchange	DOC/DOCX/ODTGIF/JPG	N/A
T4.4	Exploitation pathway: Develop and align strategic material, formats, values, guidelines on how-to for	Observational Derived - Compiled	XLSX/spreadsheets, DOC/DOCX/ODT, PDF, JPEG	N/A

12 Appendix

12.1 Appendix A: Engage4BIO open research policy

This project adopts the ethics of its object of research and therewith an ambitious open science policy. It aims at transparency (public visibility of project advancement), participation (ability for any interested individual to engage in the project) and reusability of knowledge (use of non-exclusive intellectual property schemes). These aspects are established as leading working principle in the project: by default, generated contents are thought as open and are only closed in case ethical or strategic reasons requires it. Implementing these principles allows maximizing the project impact by deepening the potential involvement of the targeted public, getting valuable feedback from practitioners and extending our online presence.

12.1.1 Transparency

Transparency is achieved by an early release policy commonly adopted in open source projects. It implements the working principles of OSS popularized under the phrase “Release early. Release often.” (Raymond, 1999). This implies sharing working documents before their completion and opening them for feedback from external stakeholders. All documents that are relevant for a wider public outside academia (such as best-practice guides, standards and paper-based tools) are written online and released under open source licenses. The early release policy is concretized in a workout loud principle adopted throughout the project. This means frequent dissemination of the performed work in order to maximize impact. This ongoing dissemination of insights and learning is crucial to encourage more companies to adopt the results in the long term.

12.1.2 Participation

Participation is achieved by the adoption of a transdisciplinary research approach (as defined by Schäfer, 2013) as described above. Documents that are relevant for a wider public outside academia are not only written online but also made open for the contribution of external stakeholders, including SMEs, open source development communities and other potentially interested stakeholder. The objective of this policy is to implement an inclusive approach to research and to maximize impact. Especially by linking the research project back to the SMEs involved. We ensure the practical relevance of the research results on the one hand, and on the other hand, find ourselves in a direct feedback loop for user-centred research. By communicating our content directly to SME in workshops, we get feedback on which parts are understandable and what needs to be revised. This also applies to the question of which parts are relevant and can be easily communicated to other SMEs. Furthermore, the interaction with the SMEs itself can also be used for impressive documentation and further communication with SMEs, e.g. as short videos or podcasts. The companies involved play the role of a testimonial.

12.1.3 Reusability of knowledge

Reusability of knowledge is ensured through the adoption of an Open Access (OA) policy for the publication of scientific works. Academic papers are published in high impact and gold open access journals and conferences as long as these exist and provide a

communication channel of sufficient quality for the content conveyed in the publication. If not, papers are published in OA using a green route. As far as it is relevant, academic results are published together with corresponding research datasets, processing software or experimentation hardware, hence following an Open Data approach. Similarly, each time it is relevant, concrete results of the project (e.g. software) are released under Open Source compatible terms.

12.2 Appendix B: Detailed Information Sheet Template

Information sheet

Dear colleague,

Your kindly invited to join us for the activities of the European project Engage4BIO and in particular for a co-creation workshop held in xxxx, on xxxx, and organised by the regional Hub xxxx, and its partners xxxx.

Before you decide to participate, we invite to read this Information sheet the project, its activities and the contribution we wish you would bring to this process.

Please, take your time to go through the information and do not hesitate to contact the project coordinator or the regional Hub coordinators in charge of this specific activity for any doubt or question (contacts are available at the end of this document).

If you wish to participate, please, we kindly ask you to read and sign the attached Consent Form and send it [to this address email xxxx, before joining the activities.](#)

Thank you for your time and contribution!

Project and objectives

Engage4BIO will strengthen circular, sustainable bioeconomy and sustainable regional development through engaging quadruple helix actors from five regional bio-based systems (hubs) in processes of design thinking, co-creation, (re)training and skills development. Engage4BIO framework is designed to be transferable, taking into account regional, natural and cultural differences and it will be applicable in different European regions.

Partners and regional hubs

The project is run by a European-wide consortium with a good balance from different EU member states. The hubs at the heart of the project build on already existing bioeconomy networks with different topics, have very strong links to industry, policy makers and associations/networks/initiatives and are well connected with other bioeconomy projects and networks - <https://www.engage4bio.eu/consortium>

The regional bio-based systems (hubs) focus on five major value chains: circular and bio-based textiles (NL); agriculture and agro-food industries (HU); wood and interior (AT); bio-based and sustainable packaging (FI); blue bioeconomy (IT). Based on their regional specificities (availability of biomass feedstock, level of technology, planetary boundaries and social capital) activities will be co-created, performed and (re-)shaped to their best advantage for each region.

The co-creation activities

The Engage4BIO co-creation activities' main purpose is to involve all relevant stakeholders in the regional areas of the 5 hubs in the co-development of 4 main outcomes to support the local uptake of innovative bioeconomic practices:

- future pathfinder manuals, collecting useful, practical and innovative ideas and instructions for activities supporting and enhancing regional bioeconomy development (vision and strategy);
- guidelines for training and mentoring for the specific value chains in the hubs, supporting the boosting of knowledge and skills useful in the bioeconomy, and in particular bio-based sectors;
- knowledge and awareness raising campaigns, based on the human-centered aspect of design thinking, involving local stakeholders, artists, designers, researchers, decision makers, professionals from the field of biotechnologies and available market actors etc. at the local level;
- good practice guidelines for local operators and innovation developers on innovative governance models supporting particularly balanced local potentials and innovation (in terms of economy, society and ecology impact) will be developed.

More information on the projects are also available in the website:

<https://www.engage4bio.eu>

Participants contribution

The participation in the co-creation activities include:

- Participation and active contribution in 1 face to face workshop of **xy** hours in **LOCATION on DATE**
- Participation in the feedback activities online (evaluation of activity, further input on purposes, confirmation of results)
- [add as relevant']
- [add as relevant']

Participants for this activity have **been chosen based on [explain briefly selection process for group composition']**.

Participants rights

Participation in the Engage4BIO activities is voluntary and free. Participants can withdraw from the project Engage4BIO at any stage. Upon request, all provided (interview) data will be destroyed and participation will not be identified at any stage of the research.

Participants data use

The Engage4BIO partner/s in charge of this activity guarantee that participants data will be held and kept confidential and used on an anonymity basis for the purpose of the co-creation activities, unless participants explicitly agree to be mentioned by name, providing explicit consent and for the following uses:

- Information and data gained through the workshop (incl. pictures, sound) published for communication and dissemination purposes;
- Publication of results of the project and of the co-creation activities to integrated in the project outputs;
- Results of the workshops shared in an open data repository.

Participants personal data and their contribution in terms of content will be stored after the end of the project for five years.

Personal data will be kept confidential and held only for the purpose of reporting and auditing as per the provisions of the Grant agreement with the European Commission.

Benefits and risks

Numerous benefits are expected for participants to engage in this activity:

- Enlarge and increase their collaboration with other local and regional stakeholders;
- Engage with other stakeholders in furthering the regional development activities and strategies;
- Increase their awareness on bioeconomy topics, services and opportunities;
- [add more as relevant related to the Hub specific sector, vision, needs, contextual aspects etc.]

No specific risks are expected from the participation in the co-creation activities, except the risk related to data protection, for which the partners have put in place already the standard preventive measures, as also explained before.

Contact and information

For further information about the project Engage4BIO and for any question and request regarding your data and for any complaint you may have about the process and handling of your participation. please, contact at any time:

- ZSI GmbH, Linke Wienzeile, 46; 1150 Vienna - email: e4B@zsi.at

For further information about this specific co-creation activity, please contact at any time:

- [Organisation name, full legal address, contact email]



12.3 Appendix C: Engage4BIO consent file for Co-Creation workshops

Declaration of Consent

I have been invited to participate in the Engage4BIO project and I have received information about the project. I understand the purpose of the project and my involvement in it.

I understand that I may withdraw from the project Engage4BIO at any stage. Upon request all provided (interview) data will be destroyed and my participation will not be identified at any stage of the research.

I understand that my data will be held and used on an anonymity basis only for the purpose of the Engage4BIO project.

I understand that, while information and data gained (incl. pictures, sound) through this interview/ survey/ co-creation workshop/ involvement may be published, it will not contain any data that could lead to my identification.

My personal information/opinion/data will remain confidential, unless I explicitly agree to be mentioned by name in the published study and/or to be quoted.

My data will be stored after the end of the project in an anonymous form for five years. My raw data will be kept confident and will only be shared in an open data repository upon my agreement.

Date:.....

Name of participant:.....

Signature of participant:

Signature of Engage4BIO representative:

For further information about the project Engage4BIO, please contact at any time:

ZSI, Linke Wienzeile, 46; 1150 Vienna - e4B@zsi.at

Please provide your contact information if we are allowed to contact you again with regard to your data (*This information will of course be stored separately from your data!*):

13 References

- European Commission-Open Access
https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm (last accessed on the 07.03.2023)
- FFG – Open Access Open Access and research data management in Horizon Europe
<https://www.ffg.at/europa/openscience> (last accessed on the 07.03.2023)
- H2020 Program – Guidelines on FAIR Data Management in Horizon 2020. By the European commission – Directorate General for Research and Innovation. Version 3.0 on 26 July 2016. URL: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf (last accessed on the 26.02.2023).
- OpenAIRE supporting open science. URL: <https://www.openaire.eu/> (last accessed on the 26.02.2020)
- Qualified Dublin Core. URL: <https://www.dublincore.org/documents/2000/07/11/dcmes-qualifiers/> <https://datacite.org/> (last accessed on the 26.02.2023).
- Raymond, E. (1999), “The cathedral and the bazaar”, Knowledge, Technology & Policy, Springer Netherlands, Vol. 12 No. 3, pp. 23 –49
- Schäfer, M. (2013), “Inter und transdisziplinäre Nachhaltigkeitsforschung Innovation durch Integration?”, Springer VS, Wiesbaden, pp. 171-194
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” Scientific Data 3 (March): 160018. <https://doi.org/10.1038/sdata.2016.18>.



“ Multi-stakeholder engagement to strengthen regional bioeconomy value-chains ”

Consortium :



ArtEZ



 www.engage4bio.eu

 info@engage4bio.eu

@Engage4BIO

