



# REFOREST

Organisation: Czech University of Life Sciences Prague

## D7.2

### Data Management Plan

Date 09.11.2022

Doc. Version 01



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 the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

**Document Control Information**

Settings	Value
<b>Deliverable Title</b>	Data Management Plan
<b>Work Package Title</b>	Project Management
<b>Deliverable number</b>	D7.2
<b>Description</b>	Data Management Plan
<b>Lead Beneficiary</b>	CZU
<b>Lead Author</b>	CZU - Martin Lukáč
<b>Contributor</b>	ORC – Colin Tosh
<b>Submitted by</b>	Martin Lukáč, Colin Tosh, Eva Májová
<b>Doc. Version (Revision number)</b>	01
<b>Sensitivity (Security):</b>	Public
<b>Date:</b>	09/11/2022

**Document Approver(s) and Reviewer(s):**

NOTE: All Approvers are required. Records of each approver must be maintained. All Reviewers in the list are considered required unless explicitly listed as Optional.

Name	Role	Action	Date
<b>Martin Lukáč</b>	Executive Board Member	Approved	24.11.2022
<b>Rym Ayadi</b>	Executive Board Member	Approved	2.12.2022
<b>Bhim Bahadur Ghaley</b>	Executive Board Member	Approved	2.12.2022
<b>Eike Lüdeling</b>	Executive Board Member	Approved	2.12.2022
<b>Julian Schwabe</b>	Executive Board Member	Approved	2.12.2022
<b>Charlotte Bickler</b>	Executive Board Member	Approved	2.12.2022
<b>Colin Tosh</b>	ORC - Researcher	Approved	18.11.2022

**Document history:**

The Document Author is authorized to make the following types of changes to the document without requiring that the document be re-approved:

- Editorial, formatting, and spelling
- Clarification

To request a change to this document, contact the Document Author or Owner.

Changes to this document are summarized in the following table in reverse chronological order (latest version first).

Revision	Date	Created by	Short Description of Changes

**Configuration Management: Document Location**

The latest version of this controlled document is stored in [Deliverables](#)

Nature of the deliverable		
<b>R</b>	Report	
<b>DEC</b>	Websites, patents, filing, etc.	
<b>DEM</b>	Demonstrator	
<b>O</b>	Other	x

Dissemination level		
<b>PU</b>	Public	x
<b>CO</b>	Confidential, only for members of the consortium (including the Commission Services)	

## ACKNOWLEDGEMENT

This report forms part of the deliverables from the REFOREST project which has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101060635. The Community is not responsible for any use that might be made of the content of this publication.

More information on the project can be found at: <http://agroreforest.eu/>

## COPYRIGHT

© All rights reserved. Reproduction and dissemination of material presented here for research, educational or other non-commercial purposes are authorised without any prior written permission from the copyright holders, provided the source is fully acknowledged. Reproduction of material for sale or other commercial purposes is prohibited.

## DISCLAIMER

The information presented here has been thoroughly researched and is believed to be accurate and correct. However, the authors cannot be held legally responsible for any errors. There are no warranties, expressed or implied, made with respect to the information provided. The authors will not be liable for any direct, indirect, special, incidental or consequential damages arising out of the use or inability to use the content of this publication.

## REFERENCE

Tosh, C.R. (2022). Task 7.5 – Data Management Plan (V1). Developed by the Horizon Europe project REFOREST ('Agroforestry at the forefront of farming sustainability in multifunctional landscapes in Europe'), funded by the European Union's Horizon Europe Research and Innovation programme under Grant Agreement Number 101060635.

## TABLE OF CONTENTS

Acknowledgement.....	4
Copyright .....	4
Disclaimer .....	4
Reference .....	4
Executive Summary .....	6
List of figures .....	7
List of Tables.....	7
List of Acronyms and Abbreviations.....	7
Data Summary .....	8
Data collection in REFOREST WP1 .....	8
Data collection in REFOREST WP2.....	8
Data collection in REFOREST WP3.....	8
Data collection in REFOREST WP4.....	9
Data collection in REFOREST WP5.....	9
Data collection in REFOREST WP6.....	10
Data collection in REFOREST WP7 .....	10
Data collection in REFOREST WP8.....	10
FAIR data .....	11
Making data findable, including provisions for metadata .....	11
Making data openly accessible .....	11
Making data interoperable.....	13
Increase data re-use.....	13
Allocation of resources .....	14
Data security.....	15
Ethical aspects.....	15
Appendix 1: References and Related Documents .....	17

## EXECUTIVE SUMMARY

The main objective of this deliverable is to implement an effective Data Management Plan and provide a data storage solution.

Data management shall oversee the balance between making data publicly accessible and respecting Intellectual Property Rights (IPR) of partners. It must adhere to the highest standards available, while ensuring the protection of third-party rights.

This deliverable sets out the key principles and operational steps to ensure data security and quality, while also fostering data exchange and cooperation. Open access to different research outputs (e.g. data, software, models, algorithms, and workflows) is provided by their deposition in trusted repositories and participation in open peer-review. All scientific publications are deposited in open access repositories upon publication. Gold articles are immediately available in the repository. All data that is made available for external use are subject to appropriate Quality Control, IPR assurance and the protection of third parties.

The Data Management Plan describes the data management life cycle for all datasets for a set number of years after the project ends.

REFOREST cloud storage solution is developed to gather the project data (i.e. qualitative and quantitative) and use them in the project work.

## LIST OF FIGURES

Figure Nr.	Title

## LIST OF TABLES

Table Nr.	Title

## LIST OF ACRONYMS AND ABBREVIATIONS

Abbreviation	Definition
DMP	Data Management Plan
WP	Work Package
AF	Agroforestry
DOI	Digital Object Identifier
KOS	Knowledge Organization Systems
IPR	Intellectual Property Rights
CZU	Czech University of Life Sciences Prague
EP	Europroject OOD
ORC	The Organic Research Centre

## DATA SUMMARY

### DATA COLLECTION IN REFOREST WP1

---

Data collection in REFOREST WP1 occurs while undertaking co-creator interviews and consultations, developing an electronic co-creation platform, developing a co-creator database, and disseminating findings. The ultimate objective is developing and supporting the co-creation process underpinning the project.

WP1 will generate: HTML code, database spreadsheets (.xmls / .accdb), and textual reports (.doc), of <500MB in total size. Outputs are novel and generated from REFOREST.

Data originate largely from interactions with WP1 co-creators (European agroforestry stakeholders) but also in creating an online platform to facilitate the co-creation process.

The data created will be useful to those interested in assessing and promoting agroforestry as a multifunctional land use strategy: government departments, public body researchers, private research and development organisations, and investors interested in natural capital. Databases of EU agroforestry stakeholders may also be of use to farmers wishing to find experts to help them develop expertise in agroforestry.

### DATA COLLECTION IN REFOREST WP2

---

Data collection in REFOREST WP2 occurs in the establishment of a knowledge inventory, the development of a living lab network, and the development of tools for AF design and performance monitoring to translate the actions undertaken in WP1 into systems that will facilitate the uptake and management of agroforestry systems by individual farmers.

WP2 will generate: textual / image inventories (reports), and HTML code, of <500MB in total size. Outputs are from previous EU projects (GFORWARD, SustainFARM, FarmTreeTool, AFINET) and novel data generated from REFOREST.

Data originate from interactions with living lab participants (European agroforestry stakeholders) and in creating an electronic platform for agroforestry design and performance monitoring.

The data created will be useful to those interested in assessing and promoting agroforestry as a multifunctional land use strategy: government departments, public body researchers, private research and development organisations, and investors interested in natural capital. Information on living lab members and developed software will also be useful to farmers wishing to find experts to help them develop expertise in agroforestry and optimise the existing agroforestry system.

### DATA COLLECTION IN REFOREST WP3

---

Data collection in REFOREST WP3 occurs while undertaking of value chain analysis, ecosystem service provision analysis, economic evaluation of agroforestry goods, environmental footprint analysis of agroforestry, valorisation of agroforestry products, and the development of economic productivity





and reruns models. The goal is to quantify the performance of existing AF systems in terms of their productivity, profitability and ecosystem service provision.

WP3 will generate text / image reports, .xmls spreadsheets, and model code, of <500MB in total size. Outputs are novel and generated from REFOREST.

Data originate from interactions with agroforestry value chain members, field ecological analyses of agroforestry systems, and original economic and environmental footprint analyses.

The data created will be useful to those interested in assessing and promoting agroforestry as a multifunctional land use strategy: government departments, public body researchers, private research and development organisations, and investors interested in natural capital. Farmers typically consider agroforestry lacking in a sound economic framework, so economic analyses will interest farmers themselves in taking up agroforestry.

#### DATA COLLECTION IN REFOREST WP4

---

Data collection in REFOREST WP4 occurs in the collation of existing datasets on landscape biodiversity and carbon sequestrations, the remote imaging of biodiversity and carbon sequestration sites, the integration of datasets and images using deep learning, the resultant development of an electronic monitoring tool for agroforestry sites, and farmers testing of this tool, to develop a monitoring and verification capability acceptable to key stakeholders in the AF value chain.

WP4 will generate: actual or linked databases on European biodiversity and carbon capture, multiple remote images of study sites, model code, and HTML code, of <10GB in total size. Outputs are from numerous sources, including previous EU projects (collated in T2.1 and 3), novel and generated from REFOREST.

Data originate from existing biodiversity and carbon sequestration databases, remote imaging, neural network model, and online tools development.

The data created will be useful to those interested in assessing and promoting agroforestry as a multifunctional land use strategy: government departments, public body researchers, private research and development organisations, and investors interested in natural capital. The ability to determine ecosystem service value of farm sites from remote images is likely to play a key role in subsidy allocation in the future, so the data will be of great interest and utility to farmers then selves and subsidy provisioning bodies.

#### DATA COLLECTION IN REFOREST WP5

---

Data collection in REFOREST WP5 occurs in the mapping and monitoring of policies and funding schemes applicable to AF, scoping of a financial model for the AF sector, application of finance models within the context of living labs, and the development of policy recommendations. The aim is to develop a financial model specific to AF and to design effective policy support.

WP5 will generate: financial models in code or spreadsheet form, a database, financial and policy reports, of <10GB in total size. Outputs are novel and generated from REFOREST.



Data originate largely through interactions with living lab members in the development of various economic tool and economic and policy analyses.

The data created will be useful to those interested in assessing and promoting agroforestry as a multifunctional land use strategy: government departments, public body researchers, private research and development organisations, and investors interested in natural capital. Findings will be of particular interest of financial and business institutions interested in harnessing natural capital, however, farmers and landowners are key developers of carbon and biodiversity "credits", so data will be of interest to those "on the ground" too.

## DATA COLLECTION IN REFOREST WP6

---

Data collection in REFOREST WP6 occurs in the development of a conceptual benchmark model, developing a develop dynamic management tool, identifying institutional barriers to agroforestry, and creating a new value chain model for agroforestry, to develop mechanisms for dynamic agroforestry system design optimisation at farm level and to develop and road-road test innovative business models for a value chain and support consultancy specific to agroforestry.

WP6 will generate: graphics, model code, HTML code, and text / image reports, of <10GB in total size. Outputs are novel and generated from REFOREST.

Data originate largely through interactions with European agroforestry stakeholders and the use of resultant data in developing business models.

The data created will be useful to those interested in assessing and promoting agroforestry as a multifunctional land use strategy: government departments, public body researchers, private research and development organisations, and investors interested in natural capital. Findings will be of particular interest to businesses involved in selling agroforestry products.

## DATA COLLECTION IN REFOREST WP7

---

Data collection in REFOREST WP7 occurs in the day-to-day project administrative and financial management, advisory board activities, production of a quality plan, during data management, and in the development of an ethical framework, in establishing the project's technical management.

WP7 will generate: text / image reports, and a graphical plan, of <500MB in total size. Outputs are novel and generated from REFOREST.

Data originate largely through interactions between researchers and administrators working on the grant and the grant's advisory board.

The data created will be useful to researchers and administrators working on the grant and Horizon Europe administrators responsible for ensuring correct procedure and value for money on the project.

## DATA COLLECTION IN REFOREST WP8

---



No new data collection, only reviewing existing data will occur in REFOREST WP8

WP8 will generate: text / image reports of <500MB in total size. Outputs are novel and generated from REFOREST.

The data reviewed are needed to ensure compliance with the ethics requirements set out in the Grant Agreement.

## FAIR DATA

### MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

---

On deposition, all data will be allocated a permanent DOI via DataCite.

Data naming within the database will follow project structure: Project identifier – REFOREST, Associated WP - WP1-WP7, File description – Self-sustaining innovation network, establishment of a knowledge inventory, etc., Date – daymonthyear, Version - v0.1 etc., Generating organisation - partner acronym (contact initials optional). E.g., REFOREST\_WP2\_Knowledge\_Inventory\_1June2023\_V1.1\_CRT.xlsx.

A maximum of five relevant keywords describing data will also be associated with each deposition to facilitate finding and re-use.

In the open access database chosen to deposit REFOREST data, Zenodo, all metadata is stored internally in JSON-format according to a defined JSON schema. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema (according to the OpenAIRE Guidelines).

To make data findable in Zenodo, metadata are assigned a globally unique and persistent identifier. A DOI is issued to every published record on Zenodo.

Data are described with rich metadata and Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments. Metadata clearly and explicitly includes the identifier of the data it describes, the DOI is a top-level and mandatory field in the metadata of each record. Metadata are registered or indexed in a searchable resource, and the metadata of each record is indexed and searchable directly in Zenodo's search engine immediately after publishing.

The metadata of each record is sent to DataCite servers during DOI registration and indexed there.

### MAKING DATA OPENLY ACCESSIBLE

---

There are three main points where the data will be deposited:

- 1) Working files will be deposited in an MS Teams area available only to project partners
- 2) "Finished product" outputs will be made available to the public through a project website
- 3) Project generated data, when suitable for release, will be placed in a project-dedicated ZENODO database available to individuals with a database account and appropriate permissions.

It is envisaged that the datasets resulting from project activities will be of an open nature, i.e., data which is freely accessible and protected by minimally restrictive or unrestricted licenses. Due to high



data demands of WP4, it may be necessary to utilise data sets from individuals or organisations not offering open access. In this case, datasets will be represented in the Zenodo REFOREST database with the email address and details of the data curator.

Generally, use of the Open Data Commons Open Database License<sup>4</sup> (ODbL) to open datasets will be encouraged, promoting the three core requirements of: attribution, share-alike and the retention of its open nature. Additional usage and sharing restrictions on the dataset will be defined through additional licenses or modifications of existing alternatives. As an alternative, the various Creative Commons licenses could be used as a licensing schema of the REFOREST processed data and also for datasets, publications, research papers and outcomes (Zervas, 2017, Karley, 2020).

Data authors may select the license that best fits their needs from the following open data licensing options (Zervas, 2017, Karley, 2020):

- Open Data Commons Attribution License
- Creative Commons CC-Zero Waiver
- Open Data Commons Public Domain Dedication and License
- Especially for public deliverables and publications, applying for a cc-by-4.0 creative commons license is suggested.

Where possible, project datasets will be released under a licence that allows them to be re-used and new work to be derived from them. Most deliverables in REFOREST output software and models, and it is quite possible that these can benefit from additional modification from individuals out with the project. Such usage is expected to increase the impact of the project beyond its duration.

Consequently, wherever possible, the 'CC-BY-NC-SA' attribution using the Creative Commons Licence Scheme will be applied. This will enable REFOREST to re-use the data and derive new work from the data and will ensure that third-party users must attribute the original dataset owner and use the same licence.

No specific software tool beyond a mainstream web browser is needed to access REFOREST data, either through the project's website or the ZENODO database.

As far as possible, final verified datasets produced by the project will be made available as open access via the project's ZENODO database. Initial versions of the datasets will have restricted access to consortium partners via the dedicated MS Teams platform. In the interests of personal data protection, restrictions will be considered for certain datasets (e.g. Socioeconomic analysis of farm operations), in which case anonymised and/or summarised findings will be made available via an openly-accessible report.

At present, there is no need for a data access committee. However, in case the datasets include personal data, access to anonymised datasets will be determined by the ethics advisor appointed to the project. Technically, any datasets with limited access can be made available through the ZENODO database only to users with an account featuring the appropriate access rights.

Data access will vary depending on storage location, and measures will be taken to enable third parties to access, re-use, analyse, exploit and disseminate the data (bound by the license specifications). In order to facilitate the interpretation of a dataset and associated third-party agreements, even in a machine-readable manner, the consortium strongly considers publishing a DCAT-AP representation for each on the project's data catalogue. Different access procedures will be implemented, enabling the export of an entire dataset and the provision of a querying interface for retrieving relevant subsets. Access mechanisms will also be supported as much as possible by metadata enabling search engines and other automated processes to access the data using standard Web mechanisms.

Registration on the REFOREST ZENODO database is a controlled process. Users may register for an account, but a definition is needed from the portal administrator. The administrator has access to the registered user's email address and name; they may verify offline the information and that the request comes from the appropriate source.

## MAKING DATA INTEROPERABLE

---

According to the Guidelines on FAIR Data Management, an EU Horizon Europe Data Management Plan is expected to address several issues relating to the interoperability of data:

- 1) Are the data produced in the project interoperable
- 2) What data and metadata vocabularies, standards or methodologies will be used?
- 3) will standard vocabularies be used
- 4) procedures in the event of unavoidable use of specific ontologies or vocabularies

Knowledge Organization Systems (KOS) are vocabularies used for the classification of re-sources based on specific topics and are available in various formats such as authority lists, classification systems, thesauri, topic maps, ontologies etc. The use of KOS ensures that the resources available within a system are classified properly and, therefore, their retrieval is facilitated. Using commonly used KOS ensures semantic interoperability allows the linking of resources available in different systems but referring to the same topic (Zervas, 2017, Karley, 2020).

In order for data to be FAIR the consortium will strive to comply or re-use existing standards whenever possible. Although original data sources may conform to different formats and standards, data processed by REFOREST will likely have been transformed into formats complying with well-known standards for the agri-food sector.

As an example, relevant standards could be (Zervas, 2017, Karley, 2020): 1) AGROVOC9 (<http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>); a controlled vocabulary for describing food, nutrition, agricultural, marine, forestry, environmental information, 2) GeoNames (<http://www.geonames.org/>); a geographical database is containing over 10 million geographical names, 3) ISO 3166-116 (<https://www.iso.org/iso-3166-country-codes.html>); the International Standard for country codes and codes for their subdivisions, 4) Ontology of units of Measure (OM) (<http://www.wurvoc.org/vocabularies/om-1.6/>), and 5) the Climate and forecast ontology (<https://www.w3.org/2005/Incubator/ssn/ssnx/cf/cf-feature>).

The consortium will re-use conceptualisations and adopt broader standards where possible (dcterms, foaf, etc.). As the project will support a Linked Data approach, when applicable, the vast majority of resulting datasets are expected to comply with semantic standards (RDF/S), and additional standardisation activities done by the World Wide Web consortium (W3C), such as OAI-ORE's JSON-LD implementation (Zervas, 2017, Karley, 2020).

## INCREASE DATA RE-USE

---

IPR holders for particular data sets will be expected to agree on the most appropriate license for their project output, however, it is expected that all project partners will work to make data freely accessible and protected by minimally restrictive or unrestricted licenses. Possible examples include Open Data Commons Attribution License, Creative Commons CC-Zero Waiver, Open Data Commons Public Domain Dedication and License, the cc-by-4.0 creative commons license.

As indicated in the "Accessibility" section above, it is expected that the vast majority of outputs will be reusable and modifiable using ShareAlike principles. We will encourage the re-use and modification of data for non-commercial and commercial use (for example, using models developed as plugins in existing electronic decision-making tools), but the commercial use of data will be done on a non-

proprietary basis. We do place some restrictions on the re-use and modification of reports and graphics to maintain the integrity of meaning as the authors intended in perpetuity.

As much as possible, we will encourage partners to make data sets available for re-use at the early possible in the project, but data sets that are fundamental to the production of scientific publications (and other outputs with similar restrictions) may need to remain embargoed until such document are published. To give all partners time to complete the publications, such data sets will be embargoed for 12 months after the completion of the project.

The licence that accompanies each dataset determines its exploitation. No significant restrictions are envisaged (currently known) for public data. Since the majority of data integrated and generated within the REFOREST catalogue will abide by the Linked Open Data (LOD) principles, the consortium will follow the best practices for supporting the life cycle of LOD. This includes its curation, repair and evolution, thus also increasing the likelihood that machine-readable structured datasets (and associated metadata) resulting from project efforts can also be of long-term use for third parties. Data quality assurance in REFOREST relates to meta-data collection, data collection, data processing and data verification.

As Karley (2020) states, tools will be prepared to optimise data quality in REFOREST. In particular, standardised templates will be developed for the collection of meta-data and data in living labs D2.1. Explicit protocols will also be written for the development of decision support and monitoring tools (D3.1, 3.2, 4.1,4.2), financial (D5.1) and policy analyses (D5.2).

Data verification will be performed on each dataset should at the point of submission to the data e-infrastructure, using: spot checks on raw data files or hard copies for outliers and errors, spot checks on data entered into electronic format for outliers and typographical errors, and proof reading of print outs of digitised data against raw data files/hard copies.

We aim to ensure data will remain available in the format they are published for a minimum of five years after the end of the project. Legacy plans for long-term preservation and curation of datasets and derived outputs will be described in an updated version of the DMP.

## ALLOCATION OF RESOURCES

---

There are various costs to be taken into account when discussing the production of REFOREST FAIR data:

- Cost of data production: these mainly include collecting and comparing data. Regarding REFOREST datasets, these costs are covered by the personnel costs estimated by partners when budgeting for tasks creating data.
- Data publication costs: these include the cost of choosing, organising and publishing the data. In REFOREST these costs are related to human activity and are covered by the personnel costs estimated by partners when budgeting for tasks. These also include the cost of publication through the Gold route in open-access journals, which is covered in 'other goods and services' of scientific partners, including the coordinator, estimated when budgeting for tasks creating data.
- Support service costs: Deposition of data into the Zenodo database is the responsibility of partners generating publicly relevant data and as such is factored into personnel time costs estimated by partners when budgeting for tasks creating data with guidance on file and metadata labelling provided by the project data manager. Creation of the project's website





and data depositions and sharing via this route are largely the responsibility of the Europroject (EP) and time and other budgetary allocations dedicated to this activity were estimated when budgeting this activity. Day-to-day data sharing via the MS Teams portal is hosted by CZU who budgeted this task when proposal preparation.

Dr Colin Tosh (ORC) will be principally responsible for data management and quality assurance on REFOREST. Dr Tosh is a computational biologist with experience of managing and ensuring public access to terabytes worth of data funded by UKRI projects of up to £1million. Data management time has been factored into ORC personnel costs on the project.

Long-term preservation of data generated by REFOREST is an issue under discussion at present. No specific funds have been requested for the long-term preservation of data beyond EU projects' normal data life cycle. Additional funds will be sought if the data generated is of exceptional value (i.e. accessed over long periods in volume) and requires long-term upkeep or updating.

## DATA SECURITY

---

Data on REFOREST's data catalogue can be made either public or private. Users who upload data can only do so in a private manner. Editors of the REFOREST's data catalogue can then review the proposed uploaded datasets and, according to their licensing and the project's policy, choose whether to make them public (hence available to visitors for downloading/querying) or not.

Partners generating data are also responsible for curating copies of their data on their secure institutional servers and the project MS Teams site (hosted by CZU), which follow the strategies for data security described above.

Sensitive (personal) data will be stored securely using an internal password-protected area on the organisational server hosted by the beneficiary partner leading the relevant analysis and these data will not be transferred.

Data storage in the Zenodo repository is expected, and it will be considered as part of the project legacy planning, including review of data security standards for long-term preservation.

## ETHICAL ASPECTS

---

A number of deliverables in REFOREST are likely to require the collection of socioeconomic data from farmers or organisations within the food supply chain (D2.1, 3.1, 5.1). For example, the development of living labs will require information on the location, size, economic performance etc of farms and other food supply chain businesses. This data will later feed into economic and policy analyses.

All personal data collected will be stored securely using an internal password-protected area hosted by the beneficiary partner leading the relevant analysis; this is a common practice in other research projects in which consortium members are involved. Personal data will not be transferred out of the host secure storage site and will not be disclosed to unauthorised users without formal consent. Persons and institutions coordinating the creation of living labs will be nominated to act as Data Controllers to ensure the safe storage of all personal data. Outputs created from the collected data will be anonymised (e.g. using alphanumerical coding) as required by the ethical review process, so that data can be exploited in the project in anonymised form. Non-anonymised data will be retained for no longer than necessary for the purpose for which it was obtained, after which it will be destroyed, and confirmed as such to the coordinator, unless instructed otherwise.



Data collection within living labs is likely to take the form of interviews and questionnaires and all human participants, whether recruited through attendance at meetings or through other advertising routes, will be provided with full details of the purposes of the research and the research activities. The researchers leading this project task (ORC with subsidiary partners) will inform participants about the purpose of the study and supply them with a project summary that they can keep. This will include information about the likely commitment involved in taking part (e.g. time needed to complete a survey) and where to get more information about the study (e.g. project website).

Participation in the studies is voluntary, and the procedures for obtaining informed consent involve using a form to seek relevant consents that will be translated into the local language. The consent form will describe the mechanisms for collecting, storing and securing personal data (aligned with EU laws for data protection and the project data management plan) and the length of time it will be held. Participants will be informed that they have the right to withdraw without being disadvantaged in any way and that the anonymity of information provided by them is assured

Consent from individuals to be included in audio-visual recordings will be obtained prior to the start of each recording activity using an agreed project consent form. Where recording will take place at a meeting or workshop, the consent form will be provided as part of the registration form. The consent form will be translated into the local language where it is needed. The above procedures and criteria used to obtain informed consent from willing participants for research and for project communication and dissemination will be subject to ethics approvals.



## APPENDIX 1: REFERENCES AND RELATED DOCUMENTS

ID	Reference or Related Document	Source or Link/Location
I	Zervas, P. (2017) DELIVERABLE 6.1 (D38). – Draft Data Management Plan. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability')	<a href="https://cordis.europa.eu/project/id/727284">https://cordis.europa.eu/project/id/727284</a>
II	Karley, A. (2020) DELIVERABLE 6.4 (D39). – Final Data Management Plan. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability')	<a href="https://cordis.europa.eu/project/id/727284">https://cordis.europa.eu/project/id/727284</a>
III	Datasets Descriptions	<a href="https://czuvpraze.sharepoint.com/:f:/r/teams/fld-t-reforest/Sdilene%20dokumenty/General/DATA%20MANAGEMENT%20PLAN/Datasets%20Descriptions?csf=1&amp;web=1&amp;e=mzAy58">https://czuvpraze.sharepoint.com/:f:/r/teams/fld-t-reforest/Sdilene%20dokumenty/General/DATA%20MANAGEMENT%20PLAN/Datasets%20Descriptions?csf=1&amp;web=1&amp;e=mzAy58</a>