# REFOREST

Organisation: Czech University of Life Sciences Prague

# D7.4

## Update of Data Management Plan

Date 13.06.2024
Doc. Version 06

**Document Control Information**

| Settings | Value |
|---|---|
| **Deliverable Title** | Update of Data Management Plan |
| **Work Package Title** | Project Management |
| **Deliverable Number** | D7.4 |
| **Description** | Prepare and release an updated version of the DMP |
| **Lead Beneficiary** | CZU |
| **Lead Author** | CZU – Martin Lukáč |
| **Contributor** | ORC – Colin Tosh, CZU – Eva Májová |
| **Submitted by** | Eva Májová |
| **Doc. Version (Revision number)** | 06 |
| **Sensitivity (Security):** | Public |
| **Date:** | 13/06/2024 |

**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 |
|---|---|---|---|
| **Martin Lukáč** | Executive Board Member | Approved | 23/06/2024 |
| **Rym Ayadi** | Executive Board Member | Approved | 23/06/2024 |
| **Bhim Bahadur Ghaley** | Executive Board Member | Approved | 23/06/2024 |
| **Eike Lüdeling** | Executive Board Member | Approved | 23/06/2024 |
| **Markus Hassler** | Executive Board Member | Approved | 23/06/2024 |
| **Julia Cooper** | Executive Board Member | Approved | 23/06/2024 |

**Document history:**

The Document Author is authorised 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 summarised in the following table in reverse chronological order (latest version first).

| Revision | Date | Created by | Short Description of Changes |
|---|---|---|---|
| 06 | 13/06/2024 | CZU – Martin Lukáč | Minor edits |
| 05 | 12/06/2024 | ORC – Colin Tosh | Section 2.4 and section 4 updated |
| 04 | 06/06/2024 | CZU – Eva Májová | Appendix 2-5 added, Appendix 1 updated |
| 03 | 14/05/2024 | ORC – Colin Tosh | Revision 02 update – size of inventories, paragraph 2.1 |
| 02 | 25/04/2024 | CZU – Martin Lukáč | Initial version update |
| 01 | 04/04/2024 | CZU – Eva Májová | Initial version |

**Configuration Management: Document Location**

The latest version of this controlled document is stored in
https://czuvpraze.sharepoint.com/teams/fld-t-reforest/Sdilene%20dokumenty/Forms/AllItems.aspx

| Nature of the deliverable | | |
|---|---|---|
| R | Report | |
| DEC | Websites, patents, filing, etc. | |
| DEM | Demonstrator | |
| O | Other | x |

| Dissemination level | | |
|---|---|---|
| PU | Public | x |
| CO | Confidential, only for members of the consortium (including the Commission Services) | |

## ACKNOWLEDGEMENT

## COPYRIGHT

## DISCLAIMER

## EXECUTIVE SUMMARY

The main objective of this deliverable is to review and update the Data Management Plan to maintain the relevance, compliance and effectiveness of the original submission delivered on December 5th, 2022. This review aims to ensure that the data are managed effectively, ethically, securely, and following relevant policies and standards.

Data management shall balance making data publicly accessible and respecting partners' Intellectual Property Rights (IPR). It must adhere to the highest standards available while protecting third-party rights.

The deliverable outlines the key principles and operational steps to ensure data security and quality while fostering data exchange and cooperation. Open access to different research outputs (e.g. data, software, models, algorithms, workflows, posters, practice guidelines, and policy briefs) is provided by their deposition in trusted repositories and participation in open peer review. All scientific publications shall be published as 'gold' access and/or their preprints will be stored in federated European Open Science Cloud repositories. All data 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.

## LIST OF FIGURES

| Figure No. | 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 Organisation Systems |
| **IPR** | Intellectual Property Rights |
| **CZU** | Czech University of Life Sciences Prague |
| **EP** | Europroject OOD |
| **ORC** | The Organic Research Centre |
| **LL** | Living Lab |
| **DPO** | Data Protection Officer |

# 1. DATA SUMMARY

## 1.1 DATA COLLECTION IN WP1

Data collection in WP1 occurs while undertaking co-creation interviews and consultations, developing an electronic co-creation and engagement platform, developing a co-creation and engagement database and innovation network, and communicating 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), textual reports (.doc / .pdf), and image inventories of <1GB in total size. Outputs are novel and generated from ReForest.

Data originate largely from interactions with WP1 co-creators (European agroforestry stakeholders and other stakeholders interested in agroforestry, initiatives, different organisations, and Living Labs) but also from established online co-creation and engagement platforms to facilitate the co-creation and engagement 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 useful to farmers wishing to find experts to help them develop expertise in agroforestry.

## 1.2 DATA COLLECTION IN WP2

Data collection in 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 HTML code, database spreadsheets (.xmls / .accdb), textual reports (.doc / .pdf), and image inventories of <500MB in total size. Outputs are from previous EU projects (AGFORWARD, SustainFARM, FarmTreeTool, AFINET) and novel data generated from ReForest.

Data originate from interactions with living lab participants (European agroforestry stakeholders) and the creation of an online co-creation and engagement platform within WP1 for agroforestry design and performance monitoring.

The created and existing data 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.

## 1.3 Data collection in WP3

Data collection in WP3 occurs while undertaking value chain analysis, ecosystem service provision analysis, economic evaluation of agroforestry goods, environmental footprint analysis of agroforestry, valorisation of agroforestry products, and developing 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 database spreadsheets (.xmls / .accdb), textual reports (.doc / .pdf), image inventories, and model codes of <1GB in total size. Outputs are from the national government institutions, and novel data are 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 created and existing data 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.

## 1.4 Data collection in WP4

Data collection in 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, HTML code, textual reports (.doc / .pdf), and image inventories of <10GB in total size. Outputs are from numerous sources, including previous EU projects (collated in T2.1 and 3), which are novel and generated from ReForest.

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

The created and existing data 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 the ecosystem service value of farm sites from remote images will likely play a key role in subsidy allocation, so the data will be of great interest and utility to farmers and subsidy provisioning bodies.

## 1.5 DATA COLLECTION IN WP5

Data collection in WP5 occurs in mapping and monitoring policies and funding schemes applicable to AF, scoping a financial model for the AF sector, applying finance models within the context of living labs, and developing 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, database spreadsheets (.xmls / .accdb), textual reports (.doc / .pdf), and financial and policy reports of <10GB in total size. Outputs are from the existing policy papers, and novel data are generated from ReForest.

Data originate largely through interactions with living lab members in developing various economic tools and economic and policy analyses.

The created and existing data 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 particularly interest 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.

## 1.6 DATA COLLECTION IN WP6

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

WP6 will generate database spreadsheets (.xmls / .accdb), textual reports (.doc / .pdf), graphics, model code, and HTML code of <10GB in total size. Outputs are from the literature research, and novel data are generated from ReForest.

Data originate largely through interactions with researchers and stakeholders (including farmers, land managers, policymakers, and others), and resultant data is used to develop business models.

The created and existing data 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.

## 1.7 DATA COLLECTION IN WP7

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

WP7 will generate database spreadsheets (.xmls / .accdb), textual reports (.doc / .pdf), and HTML code of <10GB in total size. Outputs are novel and generated from ReForest.

Data originate largely through interactions between researchers and administrators working on the project and the external ethical advisor, advisory board, and DigitAF community.

The data created will be useful to researchers and administrators working on the project to ensure correct procedures, workflow, and the best value for money and compliance with the ethics requirements in the Grant Agreement.

## 1.8 DATA COLLECTION IN WP8

No data are collected under WP8.

## 2. FAIR DATA

### 2.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

On deposition in a trusted repository, all data will be allocated a permanent DOI via DataCite.
We have instructed ReForest contributors to the database to data name the following 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 six relevant keywords describing data will also be associated with each deposition in a trusted repository to facilitate finding and reuse.

In the open-access database that stores ReForest data, Zenodo, all metadata is stored 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 complies 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.

### 2.2 MAKING DATA ACCESSIBLE

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

1) Working files will be deposited in an MS Teams space available only to project partners

2) "Finished product" outputs will be made available to the public through a project website
3) when suitable for release, project-generated data will be stored in a project-dedicated ZENODO open-access repository [Search Data Repository of the Horizon Europe REFOREST project (zenodo.org)](https://zenodo.org) available to individuals.

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, the use of the Open Data Commons Open Database License4 (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. Alternatively, the various Creative Commons licenses could be used as a licensing schema for the ReForest processed data and 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 reused and new work to be derived from them. Some of the 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 project's impact beyond its duration.

Consequently, the 'CC-BY-NC-SA' attribution using the Creative Commons Licence Scheme will be applied wherever possible. This will enable ReForest to reuse the data, derive new work from the data and ensure that third-party users must attribute the original dataset owner and use the same licence. In cases with clear commercial application for data, we will encourage the use of the CC-BY-SA 4.0 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, the final verified datasets produced by the project will be made available for 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. If the datasets include personal data, the data owner is fully responsible for anonymising datasets first unless the impacted persons provide clear personal consent that their personal data might be disclosed to 3<sup>rd</sup> party. Technically, datasets with limited access can be made available through the ZENODO database only to users with an account with the appropriate access rights. Data access will vary depending on the storage location, and measures will be taken to enable third parties to access, reuse, analyse, exploit and disseminate the data (bound by the license specifications). 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 providing 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 can access the registered user's email address and name; they may verify offline the information and that the request comes from the appropriate source.

The data will remain findable and available in the format they are published for a minimum of five years after the end of the project. Some legacy plans are currently being investigated. The current data curator, Dr Colin Tosh, has offered to curate the data in the long term, but we would ideally prefer an option that is not dependent in individuals' circumstances. We will approach EURAF (European Agroforestry Federation) to determine their willingness to curate the data in the long term.

## 2.3 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) What procedures are in place in the event of unavoidable use of specific ontologies or vocabularies?

Knowledge Organization Systems (KOS) are vocabularies that classify resources based on specific topics. They are available in various formats such as authority lists, classification systems, thesauri, topic maps, ontologies etc. Using 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 and allows linking resources available in different systems while referring to the same topic (Zervas, 2017; Karley, 2020).

For data to be FAIR, the consortium will strive to comply with or reuse 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 reuse 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).

## 2.4 INCREASE DATA REUSE

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

As indicated in the "Accessibility" section above, most outputs are expected to be reusable and adaptable using ShareAlike principles. We will encourage the reuse and modification of data for non-commercial and commercial use (for example, using models developed as plugins in existing electronic decision-making tools). Still, the commercial use of data will be done non-proprietorial. We do place some restrictions on reusing and modifying reports and graphics to maintain the integrity of meaning as the authors intended in perpetuity.

As much as possible, we encourage partners to make data sets available for reuse 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 is published (if other IPR rights do not protect the dataset). 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 most data integrated and generated within the ReForest project 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 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, processing, and verification.

As Karley (2020) states, tools are being prepared to optimise data quality in ReForest.
In particular, standardised templates were developed to collect the data in WP1 task 1.3, 1.6, WP2 task 2.2, 2.3 and WP3 task 3.1, task 3.2. An explicit protocol for the data collection in WP2 and WP3 was developed as an Addendum to MS10 - Handbook of Data Requirements and Protocols for Data Collection from the REFOREST Living Labs. Furthermore, standardised templates will be developed to collect meta-data and data in WP5 task 5.3, and standardised protocols will be developed for WP4 activities. However, WP4 is sufficiently novel and exploratory, and this will be done retrospectively as work proceeds. WP leaders are currently writing low-level workflows for 4.1 and 4.2, and this effort will intensify to produce detailed, repeatable workflows by the end of WP4 activities.
Data verification will be performed on each dataset 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 proofreading of print outs of digitised data against raw data files/hard copies.

## 3. OTHER RESEARCH OUTPUTS

No other research outputs are available or anticipated.

## 4. 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 the project 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. The 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 for this activity. Day-to-day schedule sharing via the MS Teams portal is hosted by CZU, which budgeted for this task during proposal preparation.

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

The long-term preservation of data generated by ReForest is still under clarification, as detailed in Article 2.2. No specific funds have been requested to preserve data beyond the typical life cycle of EU projects, and no additional funds will be available. The respective party will cover potential costs for the long term preservation of internal resources.

## 5. 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).

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 securely stored on each partner's secure server in compliance with each partner's Personal Data Protection Directive. Each partner is responsible for adhering to national rules and the GDPR and may not transfer these data without the written consent of the individual concerned. Where the written consent of the individual concerned is provided, these data might be stored on the MS Teams site (hosted by CZU) and be visible to other project partners. Nevertheless, these data might not be disclosed to third parties unless fully anonymised and used only for further research. The selected data storage in the Zenodo repository is part of the project legacy planning, including further review of data security standards for long-term preservation.

## 6. ETHICS

Several deliverables in ReForest require collecting socioeconomic data from farmers, other stakeholders, and/or organisations (D2.1, 3.2, 5.1). For example, developing living labs requires information on the location, size, economic performance, etc., of farms and other supply chain businesses. This data will later feed into economic and policy analyses.

All personal data collected are stored securely using an internal protected area hosted by the project partner leading the relevant analysis; this is a common practice in other research projects involving consortium members. Personal data will not be transferred from the host's secure storage site nor disclosed to unauthorised users without formal consent. Every project partner must have a DPO in place 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 to the coordinator unless instructed otherwise.

Data collection within living labs takes the form of interviews and questionnaires. All human participants, whether recruited through attendance at meetings or other advertising routes, are provided with full details of the research's purposes and activities. The project partners leading this activity inform the participants about the purpose of the activity and supply them with a project summary. This includes information about the likely commitment involved in taking part (e.g., the time needed to complete a survey) and where to get more information about the project and the project activities (e.g., the project website).

Participation in the studies is voluntary, and the procedures for obtaining informed consent involve using a form to seek relevant consent translated into the local language. The consent form describes the mechanisms for collecting, storing and securing personal data (aligned with EU laws for data protection and the project data management plan) and the time it will be held. Participants are informed that they have the right to withdraw without being disadvantaged and that none of the personal data collected throughout the activity will be shared with/ or disclosed to third parties unless fully anonymised.

Consent from individuals to be included in audio-visual recordings must be obtained before each recording activity starts using the project consent form. Where recording occurs at a meeting or workshop, the consent form is provided as part of the registration form. The consent form is translated into the local language. Participants are informed about this activity before the recording starts when participating in a recorded virtual event. Their continued presence in the meeting confirms their consent to being recorded. The above procedures and criteria used to obtain informed consent from willing participants for research, project communication, and dissemination were subject to ethics approvals.

## 7. OTHER ISSUES

Other national/funder/sectorial/departmental procedures for data management are not being / won´t be utilised.

# 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') | https://cordis.europa.eu/project/id/727284 |
| 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') | https://cordis.europa.eu/project/id/727284 |
| III | Datasets Descriptions | https://zenodo.org/communities/reforest/records?q=&l=list&p=1&s=10&sort=newest |
| IV | Addendum to MS10 - Handbook of Data Requirements and Protocols for Data Collection from the REFOREST Living Labs | https://czuvpraze.sharepoint.com/:w:/r/teams/fld-t-reforest-WP3AFsystemperformance/_layouts/15/Doc.aspx?sourcedoc=%7B2E333159-B563-4164-83C5-D3E285AD950C%7D&file=REFOREST%20LLs%20Data%20protocols_v3.0.docx&action=default&mobileredirect=true |
| V | Excel-based data collection protocol for WP3, Task 3.2 | https://czuvpraze.sharepoint.com/:x:/r/teams/fld-t-reforest-WP3AFsystemperformance/_layouts/15/Doc.aspx?sourcedoc=%7B8DA14C70-9839-4800-95B7-3D1AE7997346%7D&file=Reforest%20WP3%20Data%20collection%20sheet_5Feb_2024.xlsx&action=default&mobileredirect=true |

**REFOREST**

## Content planning

YOUR NAME AND RELEVANT LIVING LAB: .........................................................................................

| What to focus on? | EXAMPLES OF POTENTIAL FORMATS: |
|---|---|
| **A broader topic in the field of agroforestry (AF)**<br>- Methods and approaches in AF research<br>- Benefits of AF (environmental, economic, social) – concrete local examples<br>- Evolution of AF practices – current trends, working models<br><br>**AF in your LL**<br>- Advantages/disadvantages of a specific AF practice applied in the LL<br>- Effect of AF on a specific element (soil, biodiversity, etc.) – how do you prove it scientifically? What kind of measurements do you have?<br>- Crop rotation strategies – what is your experience and advice?<br>- AF as part of your strategy for climate adaptation<br>- Work progress – results, challenges, risks, findings, conclusions, next steps<br><br>**Economic viability and society**<br>- Marketing strategy you apply<br>- Do you engage external people and how?<br>- Economic constraints and impact on agroforestry design / practice / approach<br><br>**Practical tips**<br>- Practical tips for establishing and managing agroforestry systems<br><br>**Other**<br>- Curious facts – did you know that...... | o   Articles, blog posts<br>o   Interviews<br>o   Newsletter<br>o   Studies, case studies<br>o   Webinars, events<br>o   Infographics, charts<br>o   Videos, podcasts<br>o   Polls, surveys<br>o   Reports, briefs, reviews |

**REFOREST**

| | |
|---|---|
| - Quizzes - to spark interests in stakeholders, to check what is the opinion of people in the field<br>- Events you've attended and presented the project and its results<br>- Quotes, motivations, inspirational thoughts | |

**List at least 3 content materials you can provide by the end of 2024. What will be the format? When will you deliver it to EP (month and week) – please try to distribute your materials evenly in the period?**

1. 'Exploring Agroforestry in an Upland Landscape' webinar recording – could provide it in first week of June?

Agricology case study featuring living lab farm – first week of August?

3. A colleague has produced a review looking at woodland grazing to improve ecology whilst ensuring the method is economically viable – particularly linking in to one of our case study farms, Gowbarrow Hall. We want to look at building on this – and developing it into a useful output / tool, drawing out key points and exploring with living lab farmers – last week of November?

_**Good example of a work progress article:**_ _https://www.organicresearchcentre.com/news-events/news/quantifying-impacts-of-agroforestry-on-soil-health/_

**REFOREST**

## Have you attended or organized an event related to ReForest?

## Tell us the story!

Dear ReForest partner,

You are kindly invited to provide information on events and initiatives you have been involved in to represent or promote ReForest project, either as participants or organizers. The collected data will be primarily used to generate news articles for the project's website and social media, so that we popularize our activities through informing the public about what we have been doing as a project but also, where they can find us. With regard to the latter, please do not limit yourself to describing only past events, but include planned future events as well, so we can announce/advertise them on project's digital channels.

Thank you for your time in sharing this information and contributing to a good dissemination impact of ReForest and establishing it as a power of authority in its field.

For questions or further information, contact ███████████████████████

| |
|---|
| Name of event: |
| Type of event: |
| Date and location of event: |
| Link to event (if any): |
| |
| Relevant pictures (2-3 pictures possibly in the following format - 510 x 320 pixels): |

**REFOREST**

## Content for case study booklet

| Sections | Description |
|---|---|
| **Name of the farmer, farm and location**<br>*Please, add 1 photo of the farmer+2-3 high quality photos of the farm* | |
| **The farm** (200 words)<br>*Please introduce yourself and your farm;*<br>*What you farm, your farm system and your general farming approach.*<br>*Please give a brief overview of how agroforestry fits into your farming system.* | |
| **The way to agroforestry** (300 words)<br>*Why did you decide to try agroforestry as a land management practice?*<br>*What factors have influence your decisions in practicing agroforestry? (i.e neighbours' reactions, lack of knowledge, political or governmental policy / support, uncertainty of the future, short versus long-term gain etc.)*<br>*How have you acquired knowledge on agroforestry?*<br>*Are you part of any networks?*<br>*Where do you feel your knowledge gaps lie?* | |
| **Agroforestry in practice** (300 words)<br>*What are your main practical agroforestry challenges and how have you / do you overcome them?*<br>*How has or does agroforestry benefit your farm/holding (e.g. income, biodiversity, soil quality etc.)?*<br>*What practical recommendations do you have for fellow farmers / growers who may want to deploy similar practices and techniques to you?* | |
| **What is ahead?** (150 words)<br>*Please describe the future goals you have for your farm/holding, including any planned agroforestry.*<br>*What are you hoping to achieve from being involved in the living lab?* | |

## Combined research approach for achieving WP2 task 2.3 and WP3 task 3.1

To reach the respective objectives and desired results of WP 2, task 2.3 (Identifying knowledge gaps) and WP 3, Task 3.1 (Value chain analysis of AF systems) a combined research approach seems necessary because first to avoid an overload of research requests towards the living labs and second to simplify the research process. Relevant questions targeting the identification of knowledge gaps can be easily included in the context of a value chain analysis and the value chain analysis provides starting points that support knowledge gap identification.

*Table 1: Roadmap*

**Timeline of delivery**

| First 3 Weeks in March | April and Mai | June to October |
|---|---|---|
| Desk Study | Stakeholder surveys/ data collection | Data evaluation and report creation |

## Methodology

### Desk Study

The main methodology of working on this research approach will be stakeholder surveys with persons who are involved in AF systems. Before these surveys can take place, a desk study is needed. During the desk study relevant stakeholders will be identified. The starting point is hereby the depiction of the living lab descriptions in MS8_REFOREST_v4. A selection between the described living labs is required, because not every living lab represents an actual value chain participant. At the end of this paragraph a tabular preselection can be found. If one of our REFOREST partners finds their live lab excluded and thinks it has been erroneously sorted out and should be included in our value chain analysis research, please contact us. After the first stakeholders have been identified in this way, further information on them is researched. This serves to

- determine existing knowledge and thus avoid duplication in the data collection process and,
- to adapt the questionnaire individually to the respective stakeholder and thus to be able to determine information that is as specific as possible during data collection.

University of Marburg
Department of Geography
Working Group on Regional Studies

**REFOREST**

**Outreach to interview partners**

Since the living labs are spread all over Europe, there are both geographical difficulties due to distance and language barriers when collecting data. Our research team is limited in language use to English and German, for this reason we may need the support of the project partners. The extent to which help will be necessary will become apparent during the preceding desk research and may differ from living lab to living lab.

Support activities from partners may include:

- Language check of questionnaires
- Forwarding the questionnaire to LL-representatives
- Conducting interviews
- Provide feedback to follow up questions

It is also worth mentioning that while the living labs and the stakeholders connected to them provide a key target group for these research tasks, the research needn't be limited by living labs, so it is possible that new stakeholders can be identified during the desk study and stakeholder surveys. Such stakeholders (for example AF practitioners outside of the LL network) will be contacted individually.

*Table 2: Stakeholder preselection for further research towards the desk study.*

| Country and Location | Lead contact | AF stakeholder description |
| --- | --- | --- |
| Denmark, Taatstrup | ██████ ██████ | Combined food and energy production system that provides Wood chips (willow), food crops (wheat, barley, oats etc.) and grass-clover ley |
| Germany, Obermoschel | ██████ ██████ | Hof Lebensberg: Organic (BIOLAND certified) Broadleaf trees, fruits & nuts, keyline system (basic), CSA – community supported agriculture that provides vegetables, herbs and |

University of Marburg
Department of Geography
Working Group on Regional Studies

| | | |
|---|---|---|
| | | perennial specialties, cereals, potatoes, berries, fruit and nuts (local and exotic) |
| Hungary, Olaszfalu | ██████ ██████████ | Family farm: wood pasture that provides animal, food and services (horse riding, riding camp, education and training, rural festivals venue) |
| Hungary, Vertesacsa | █████ ██████████ | Family farm: alley cropping, shelterbelt and wood pasture these provide high quality handmade products from native fruit species, forage, fuel and services (education and training) |
| Bulgaria, Indje Voivoda | █████████ | Farm: forest farming and silvopastoral system these provide honey, fruits and animals |
| Czech Republik | █████ ██████ | Policymakers Farmers: mostly silvopastoral, but a bit of silvoagricultural too, these provide fruits, from the long term perspective wood and classical agricultural products like cereals/oilseeds, cattle/sheeps AF stakeholders |
| United Kingdom, Northern England | █████ █████████████ | Cockle Park Farm, RegenFarmCo and Ings Farm, these provide Arable crops (wheat, barley, beans) and wood biomass for energy |

**University of Marburg**
**Department of Geography**
**Working Group on Regional Studies**

██████████████████
███████████████████

**REFOREST**

**Stakeholder surveys**

The survey outlined below shows the information we seek through desk study and through interviews.

As already mentioned, the surveys are customized for each stakeholder, so it is currently not possible to define which type of survey is suitable and feasible for each stakeholder. The ideal survey will be held as semi-structured face to face interviews, which is an approach that keeps the conversation flexible while retaining comparability and structure. Other communication approaches, such as video-based communication via Zoom or other services, are also conceivable. Based on explicit permission of the interviewees, the conversations will be recorded. If permission is not granted, notes will be taken. Questions will, in general, be asked in an open manner in order to encourage the interviewee to elaborate. To make it easier for the interviewer to keep track of the questions and to ensure that all relevant questions are addressed during the interview, an interview guideline is used. The guideline is derived from the general survey questions below and which takes into account findings from the desk study as well as the expertise of the interview partner. Therefore, the guideline of each interview will vary.

| Questions | (potential) theoretical links |
| --- | --- |
| **1. Emergence & key farm characteristics** | |
| **When and how were the farm's operations started? / Since when is your farm practicing agroforestry? What were the main motivations for it?** | |
| • years active; origin of idea/knowledge; mission & objectives; conversion to agroforestry practices; support of other stakeholders in start-up process | GVC: Governance |
| • *Only to ask, if farm is part of a producer network or cooperating with other farms:* Could you describe your role within the production network/cooperation as well as the relationship and the arrangements with the producer network farms? How and by whom is the production network managed? | |
|     o Governance structure of network; Farm's role; number of actors involved | |
| **Coming to some key farm characteristics, could you describe your farm in terms of size, farming practices and organizational structure? Could you also outline why you have chosen this particular farm design?** | GPN: firm architecture & ownership; GVC: governance |
| • Farm size (ha); farming practices; reasons; how farmer obtained knowledge about it | |
| • Organizational structure (legal status; ownership of farm and land; responsibilities for management & risks; labour force etc.) | |

University of Marburg
Department of Geography
Working Group on Regional Studies

## 2. Value chain

*Agricultural production*

| | |
|---|---|
| **What are the products/services of your farm? How do you produce them and why?** | GPN: value creation |

- Product/service types (crops, livestock etc.); purpose; production scale; particularities of cultivation; main products in terms of quantity and economic importance? Why?

- How are you dealing with co- or waste-products? (e.g. reuse, sale, external disposal)

- Do you offer any supplementary products or services? (non-agricultural activities carried out, e.g. workshops) If yes, why and what is it about?

- *Only for network producers or cooperation farms:* Which of your products & services are devoted to the production network? Can you estimate the share of your land and of your whole production capacity/quantity devoted to this purpose?

| | |
|---|---|
| **How do you plan your annual production and what are important factors for the planning?** | GVC: governance |

- Decision making strategies; involvement of other stakeholders? (e.g. consumers)

| | |
|---|---|
| **What aspirations do you have to your products and production process and how do you achieve them? Do you follow any environmental or social production standards or schemes? Why or why not?** | GPN: value creation & enhancement |

- internal self-commitments (e.g. traceability, environmental compatibility, regionality, seasonality etc.); official standards (e.g. organic); viability of implementation?

*Inputs & supply*

| | |
|---|---|
| **What kind of supplies do you have to buy for your production and where from? Can you describe how you select suppliers and what relationship you have to them as well?** | GVC: governance; GPN: embeddedness/ power |

- Type & source of purchased supplies (e.g. seeds, fuel, feed, fertilizer); largest quantities & expenses (shares); selection criteria; length & nature (formal/informal) of relationship; number & spatial range of suppliers; arrangements with suppliers; difficulties (e.g. dependency)

- Are there any supplies that you can provide yourself or without buying it?

| | |
|---|---|
| **Could you outline how farm land, labour as well as facilities and machineries are used and organised? What are the reasons for the way you organize each?** | |

- Land use, spatial distribution and acquisition; reasons

- Number & type of workers (e.g. full-time/seasonal, family, volunteers); (relevance of) professional background of workers; training offered; division of labour; reasons

  *GPN: firm's architecture*

- farm facilities (buildings, storage, other infrastructure etc.); type of machines owned/rented/shared & used; repairs; maintenance; reasons

University of Marburg
Department of Geography
Working Group on Regional Studies

*Processing & Packaging*

### How and by whom are your products processed and packaged? Why?

- In-house or outsourced processing/packaging; in which form are products processed/packaged; material/machineries used; innovative packaging; challenges

*GVC: governance; GPN: value enhancement*

*Distribution channels, marketing & sale*

### Could you describe the distribution of your products? Please track the flow of the products until they reach the consumer. Could you also estimate the revenues, quantities and labour for each of the distribution channels?

*GPN: value capture, embeddedness; GVC: governance*

- Design, diversification and length of distribution (spatial range, number & type of actors involved); Why selling through these channels? ; To whom are you (mainly) selling?; How do you manage the logistics (transport, infrastructure; storage etc.)?; What difficulties occur regarding selling your produce (for the different channels)

- Quantities: How much of the produce is sold through each of the channels?

- Revenues: Most rewarding channel &product group (financially, personally)? Why?

- Labour: Which channel/produce involves most labour?

### How does the marketing of your products look like? What particular steps have you taken to add value to your products and how successful would you say it is?

*GPN: value enhancement*

- Marketing strategy; added-value activities (e.g. unique costumer experience; story telling; role of ICT/social media; labelling/certification; communication of farming philosophy/system & its benefits; difficulties in finding costumers?

### How are prices determined in your sale channels? To what extent do you consider the prices being fair/satisfying (for you but also for consumers)?

*GPN: value capture, power*

- Determination of prices; true costs/benefits reflected in prices

- Are there any differences in prices to be made depending on the type of costumer/sale channel/product and if yes, why?; In which of your products is the most value added for you as a producer and why?

### How would you characterize the market for your products and your consumers?

*GPN: power, value, embeddedness*

- How competitive is the market? How does this show? ; What are recent market developments for your products offered? What influences it?

- Characteristics of consumers; Why are they buying your products (selling points)?

### 3. Differentiation from conventional systems

### How does your AF-food production system differ from conventional ones and how does this impact the overall market opportunities of your farm?

- Particularities of AF-production system/products (e.g. prices, quantity, quality, costumer relation; values; knowledge sharing; labour-, time-, resource-, knowledge-intensity)

- To what extent are these particularities reflected in your prices & marketing?; How is this appreciated by your costumers (e.g. willingness-to-pay, trust)?; To what extent does this impact the competitiveness of your products?

## 4. Evolution & financial situation of farm

**If you compare the farm's position today to its initial phase, how did it change since then and why?**

GPN: firm architecture

- internal changes (mission & values, ownership, legal status, suppliers, costumers, production quantity/efficiency, labour force, diversification, cultivation technique etc.)

- changes in/triggered by the external environment (e.g. regulations, subsidies, partnerships, market entrance, changes in demand/competition etc.)

**Could you also describe the evolution of your financial situation? How do you (re-)finance your farm now and how was it in the start-up phase?**

- Profitability; timing of exceeding break-even; main expenses & income sources, why?; financial support (where from?); capital investment; access a fund/grant/subsidies/trust; farming as main occupation

GPN: power; value capture

- How quickly is your farm able to bounce back from external shocks/variable costs?

## 5. Drivers & barriers

**Which factors would you say most enabled you to establish and maintain your AF-production system and limit risks?**

GPN: embeddedness

- Drivers (e.g. regional support mechanisms; politics, skills, assets, social capital; partnerships & networks; trainings; learning from others; research projects, media etc.)
- Do you feel supported in the way you farm by politics, regulations, society etc.?

*If not already mentioned as "driver":* **How important are linkages to other actors for your farm and/or an enabling environment? What is their role exactly?**

- Type of actors (e.g. political/financial institutions; local economy/community; other farmers; interest groups etc.); actors' role (value chain, advice, information, supporting, knowledge etc.)

GPN: Embeddedness; power

- Value chain linkages: How difficult is it, to coordinate and find suitable (sub-)suppliers and (sub-)purchasers? ; How do you strike the balance between the farm's autonomy and the need to collaborate with other stakeholders?

**What were/are the main challenges that you were/are facing?**

- E.g. financial/political uncertainties; availability of local/regional infrastructure for key operations; (qualified) labour availability; access to & sufficiency of land; trade-offs between long-term benefits & short-term challenges etc.

future needs

- How did you manage to overcome these challenges?

University of Marburg
Department of Geography
Working Group on Regional Studies

**6. Future needs**

What would you need or what would need changing to further develop and scale-up your production system?    Future needs

**7. Knowledge gaps**

Could you think about if there are any knowledge gaps regarding AF practices?
- for example establishing AF and enhancing existing AF-systems

Could you think about if there are any knowledge gaps regarding AF business models/economics/investments?

How easy is it for farmers to get access to AF knowledge? Which networks provide access to AF knowledge or through which stakeholders access to AF knowledge is granted?

**Data evaluation and report creation**

Interviews will be audio-recorded and if this is not possible, notes will be taken. The audio records will be transcribed. After that the answers of different interview partners in the transcripts and notes will then be allocated in a matrix which is structured broadly based on the categories outlined in the table below (this may still be adapted).

*Table 3: Survey matrix*

| Main question | Sub-questions/Aspects to be discussed | Directed at task: |
|---|---|---|
| **Emergence and key characteristics:** When/how/why did you start using AF-practices? | - Motivation<br>- Buildup of expertise/training/knowledge input<br>- Investment/financing<br>- Establishing marketing channels<br>- Support from external stakeholders | WP3, task 3.1 |
| **Value Chain** about the farm and its AF-characteristics: | - Products<br>- AF practices (plants, practices, technologies)<br>- Processing | WP3, task 3.1 |

University of Marburg
Department of Geography
Working Group on Regional Studies

| | | |
|---|---|---|
| | - Marketing channels<br>- Customer characteristics<br>- Input suppliers<br>- Costs, revenue, prices | |
| **Competition:** How does the AF practitioner view competition from conventional farming? | - Cost, work effort, revenues<br>- Regulations | WP3, task 3.1 |
| **Barriers:** What are current main problems/barriers for engaging in AF practices?<br><br>**Needs:** What would need to change to encourage conventional farmers to introduce AF practices?<br><br>**Knowledge gaps:** Which knowledge gaps needs to be closed to support or implement AF practices? | - Regulations<br>- Expertise/training/knowledge<br>- Needed inputs<br>- Competition from conventional practices<br>- Openness towards new practices<br>- Knowledgebase about AF practices<br>- Knowledgebase about AF business models<br>- Accessing AF knowledge | WP2, task 2.3, WP3, task 3.1 |

Insights from interviews may be complemented by field visits as well as fact/figures and other data provided by documents or publicly accessible information. The general aim is to reconstruct AF systems in their make-up (business model, input-output structures) within a report, to identify the socio-economic frameworks in which they can emerge and to identify the main barriers/knowledge gaps that farmers prevent engaging in AF practices.

University of Marburg
Department of Geography
Working Group on Regional Studies