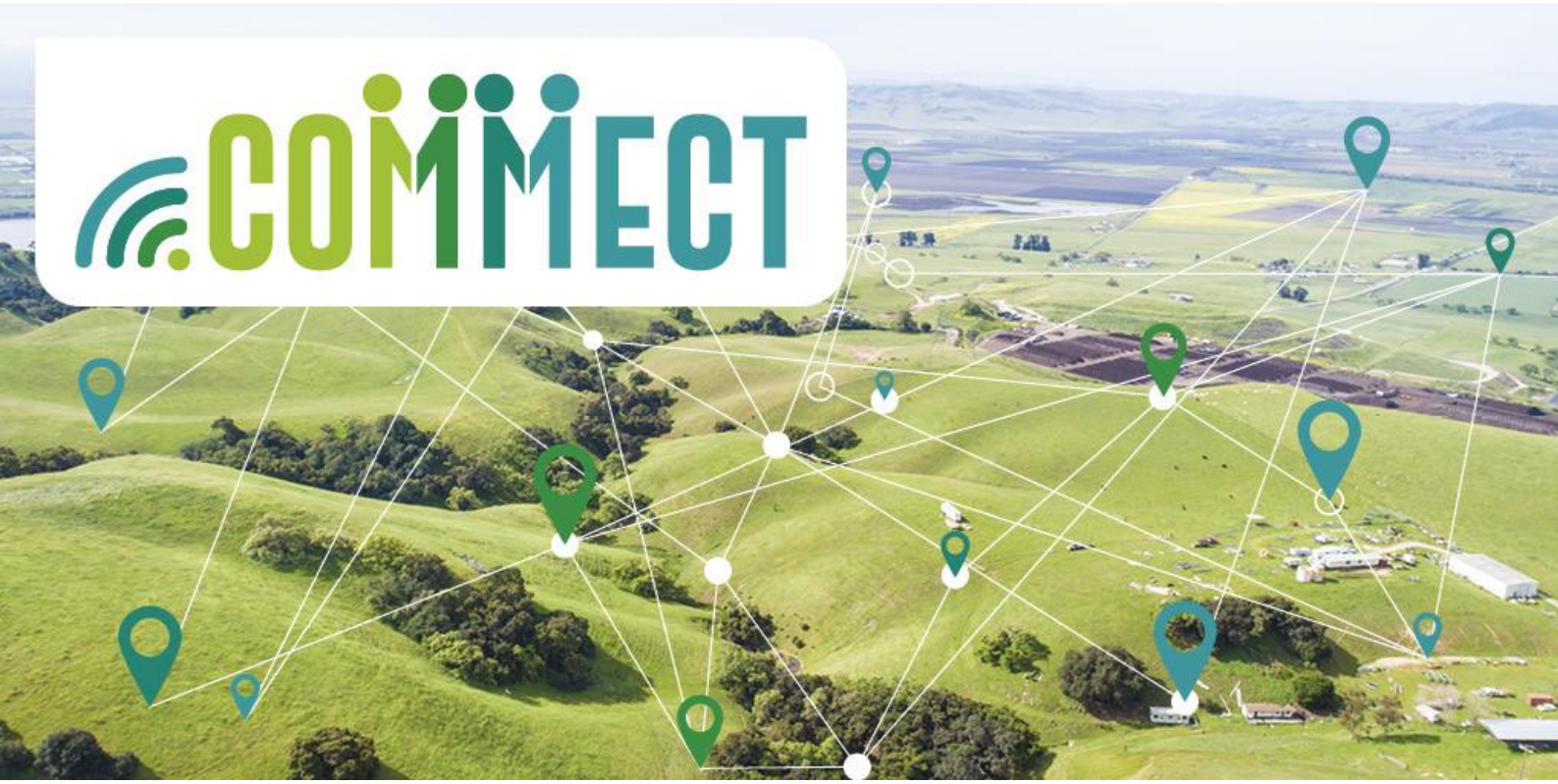
Bridging the digital divide and addressing the need of Rural Communities with Cost-effective and Environmental-Friendly Connectivity Solutions



Deliverable 7.2
Data Management Plan – Version 1

February 2023

PUBLIC

This project has received funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement no. 101060881.

COMMECT

Bridging the digital divide and addressing the need of Rural Communities with Cost-effective and Environmental-Friendly Connectivity Solutions

Grant agreement No. 101060881

Data Management Plan – Version 1

WP7 Project Management

Deliverable ID	**D7.2**
Deliverable Name	**Data Management Plan – Version 1**
Lead Partner	Aalborg University (AAU)
Contributors	Luxembourg Institute of Science and Technology (LIST), Telenor (TNOR), Turkcell (TCELL), Dunavnet (DNET)

PUBLIC

COMMECT Project Abstract

Over the last years, the importance and need for broadband and high-speed connectivity has constantly increased. The Covid-19 pandemic has even accelerated this process towards a more connected society. But this holds mainly true for urban communities. In Europe a 13% lack access persists, and mainly concerns the most rural and remote areas. Those are the most challenging to address since they are the least commercially attractive. COMMECT aims at **bridging the digital divide**, by providing quality, reliable, and secure access for all in rural and remote areas. The **goal of extending broadband connectivity in rural and remote areas** will be achieved by *integrating Non-Terrestrial Networks with terrestrial cellular XG networks, and low-cost Internet of Things (IoT). Artificial Intelligence, Edge and Network Automation will reduce energy consumption both at connectivity and computing level.*

Participatory approach with end-users and ICT experts working together on development challenges will be the key **for the digitalization of the sector**. To ensure the rich exchange of best-practice and technical knowledge among the actors of the agro-forest value chain, COMMECT will set up **five Living Labs across and outside Europe**, *where end-users "pain" and (connectivity) "gains" will be largely discussed, from different perspectives.*

COMMECT aims at contributing to a balanced territorial development of the EU's rural areas and their communities by making smart agriculture and forest services accessible to all. COMMECT will facilitate that, by developing a **decision-making support tool** able to advise on the best connectivity solution, according to technical, socio-economic, and environmental considerations. This tool, incorporating collaborative business models, will be a *key enabler for jobs, business, and investment in rural areas, as well as for improving the quality of life in areas such as healthcare, education, e-government, among others.*

Document Information

Document Name	Data Management Plan
Version No.	V1.0
Due Date Annex I	28/02/2023
Report Date	27/02/2023
Number of pages	20
Lead Authors	Melisa López (AAU)
Other Authors	LIST, TNOR, TCELL, DNET
Dissemination Level	Public

Document History

Version	Date	Description	Authors	Reviewed by
0.0	13.01.2023	Document Draft Creation	M. López	WP Leaders
0.1	07.02.2023	Complete Draft	M. López	Richard O. Agyemang (TNO) Patrik Hitzberger (LIST)
1.0	14.02.2025	Minor Reviews	M. López	M. R. Palattella (LIST)

Document Approval

Version	Name	Role in the Project	Beneficiary	Date
1.0	M.R. Palattella	Project Coordinator	LIST	24.02.2023

Executive Summary

The Data Management Plan (DMP) contains the measures that the Horizon Europe (HE) COMMECT project will take to follow the guidelines of the European Union (EU) for open access to scientific knowledge produced within the European funded projects [1]. The members of the COMMECT consortium will establish through this document mechanisms for allowing open access to the project results. These will be in the form of project deliverables, scientific publications, collected data from the five different Living Labs (LLs) and availability of open-source software contributions (e.g., Decision-making Support Tool).

The COMMECT DMP follows the structure of the Horizon Europe DMP template [2]. It includes the guidelines that will be followed along the project for quality assurance of all the collected data and generated documentation. It describes what data will be collected, processed, or generated, what methodologies and standards will be followed, whether and how this data will be shared and/or made open, and how it will be curated and preserved. It also ensures that the project complies with the applicable data protection regulations, including the European General Data Protection Regulation (GDPR). For data collected in the LLs, the DMP outlines data and metadata formats, storage, sharing, and data security methodologies.

The information included in this document provides an insight of how the COMMECT project will comply with the Findable, Accessible, Interoperable, and Reusable (FAIR) principles for Open Science established by the EU.

Table of Contents

COMMECT Project Abstract **3**
Executive Summary **5**
Glossary of Terms **7**
1. Introduction **8**
2. Data Summary **8**
3. Data Management **10**
3.1. Internal Data Management 10
3.1.1. LL N.1 - Luxembourg – Digitalisation of Viticulture 11
3.1.2. LL N. 2 - Norway – Connected Forestry 11
3.1.3. LL N. 3 - Denmark – Connected Livestock Transport 12
3.1.4. LL N. 4 - Türkiye – Smart Olive Tree Farming 12
3.1.5. LL N. 5 - Serbia – Sustainable Agriculture and Preservation of Natural Environment 13
3.2. Open Access to Research Data 13
3.2.1. Findability 14
Metadata *14*
3.2.2. Accessibility 15
3.2.3. Interoperability 16
3.2.4. Reusability 16
4. Data Security **17**
4.1. Internal Data Security 17
4.2. External Data Security 17
5. Other Research Outputs **17**
5.1. Deliverables 17
5.2. Scientific Publications 17
6. Allocation of Resources **17**
7. Data Management Responsible **18**
7.1. Data Responsible 18
8. Conclusion **18**
Appendix 1: Data Responsible **19**
References **20**

Glossary of Terms

AI	Artificial Intelligence
API	Application Programming Interface
DMP	Data Management Plan
DOI	Digital Object Identifier
DST	Decision-making Support Tool
EO	Earth Observation
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable
GDPR	General Data Protection Regulation
HE	Horizon Europe
IoT	Internet of Things
KPI	Key Performance Indicator
LCA	Life Cycle Assessment
LL	Living Lab
NTN	Non-Terrestrial Networks
OA	Open Access
RAN	Radio Access Network
WP	Work Package

1. Introduction

The COMMECT project aims at bridging the digital divide, by providing quality, reliable, and secure access for all in rural and remote areas. It will leverage existing and evolving access technologies: 2G/3G/4G/5G and Non-Terrestrial Networks (NTN) to ensure reachability everywhere. Access to fast broadband Internet will enable the mainstreaming of smart farming, forestry inventory and monitor changes, including small and big Internet of Things (IoT) data, Earth Observation (EO) data; and allow the application of Artificial Intelligence (AI) to solve business and social issues for the communities.

The project will provide the basis of the connectivity solutions for rural areas and communities by carefully evaluating the heterogeneous needs of the end-users in various countries, regions, and sectors. Socio-economic and environmental factors will be key in the design of the COMMECT connectivity platform to increase the sustainability, competitiveness, and resilience of those communities. Through the work performed in the different Work Packages (WPs) of the project, COMMECT will open new business opportunities and enhance the capability of rural communities to set-up innovative, collaboration- and service-focused business models.

To ensure the rich exchange of best practices and technical knowledge among the actors of the agro-forest value chain, COMMECT will set up five Living Labs (LLs) across and outside Europe, where end-users “pain” and connectivity “gains” will be discussed from different perspectives. User needs will be collected in WP1 and used to define a series of use cases. Using knowledge exchange (best-practice and technical) and lessons learned from LL activities in WP4, COMMECT will design and evaluate (from a technical, socio-economic, and environmental perspective) in WP2, WP3, and WP5 the solutions for meeting the needs of end users. Additionally, COMMECT also develop in WP3 a Decision-making Support Tool (DST), supporting decision-making for several actors in the agricultural industries.

The activities abovementioned and the work performed during the project are expected to generate different results in the form of deliverables, scientific publications and experimental data collected during multiple measurement campaigns in the LLs. The consortium plans to make most of these results publicly available, following the Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon Europe [1], according to the strategy and plan outlined later in this document.

2. Data Summary

As mentioned in Section 1, the project will collect end-user needs and use the collected information to establish end-user requirements, for a set of relevant use cases. Afterwards, the connectivity solutions will be designed accordingly. These solutions will not only be analysed from a technical perspective, but also from a socio-economic and environmental point of view.

Therefore, COMMECT involves the collection and processing of different types of data. For the end-user needs and the socio-economic impact assessment, personal data will be collected. This information collection will be governed by standard research ethics principles. The project will conform to current EU legislation on the collection, storage, protection, retention, and destruction of personal data. In particular, compliance to Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation) and the European Court of Justice ruling on the “Right to be forgotten” (May 2014) will be ensured. Compliance will also be maintained against individual national

legislations of the countries of origin of the COMMECT partners. COMMECT does not envisage collecting personal sensitive data. But as a matter of course the project will treat all personal data as sensitive data and it will be anonymized before being shared with other partners. For data collection involving humans, COMMECT will implement informed decision procedures, giving a subject adequate information concerning the study, providing adequate opportunity for the subject to consider all options, responding to the subject's questions, ensuring that the subject has comprehended this information, obtaining the subject's voluntary agreement to participate and, continuing to provide information as the subject or situation requires.

Considering the abovementioned, the COMMECT project distinguishes between different types of data collected along the different phases of the project:

- **End-user Data:** Opinions, problems, and needs collected from different users during the user needs collection and use cases definition stage. This data will be collected through surveys, workshops, interviews and/or meetings. It is considered critical since it may contain personal data and it needs to follow General Data Protection Regulation (GDPR), so data should be anonymized prior to sharing with other partners.
- **Agricultural Data:** Raw data generated by the different sensors and devices in the LLs. Examples of this are data collected in the fields (LLs in Luxembourg, Türkiye and Serbia), and in the forest (LL in Norway). For the connected livestock transport LL (in Denmark), collected data will not involve data from the truck loggers, and therefore we will only consider *technical, socio-economic, and environment* data types (see next items). Agricultural data will be used for developing the different applications for the LLs, as well as the DST.
- **Technical Data:** Data gathered during the technical performance assessment of the connectivity solutions designed and implemented to overcome users' needs. Data will be collected in both a controlled environment and in the LLs. Some examples of technical data could be Radio Access Network (RAN) data, wireless performance data or use case-specific data (sensor or EO data, data for remote control of machinery, etc.). For technical data, data protection and/or copyright/license restrictions might apply (e.g. RAN data), and these will be verified for all collected data sets and might prevent opening them.
- **Socio-economic Data:** Qualitative data gathered to develop an understanding of the impact of connectivity on various parts of rural life. Multiple surveys will be conducted to measure how connectivity affects perceived quality of life, schooling and community building. Considered critical since it may contain personal data and it needs to follow GDPR, so data should be anonymized prior to sharing with other partners.
- **Environmental Data:** Environmental indicators such as energy or CO2 emissions will be collected for each of the connectivity solutions. These indicators will be used to evaluate the environmental impact using a Life Cycle Assessment (LCA) framework. The gathered environmental data is not considered sensitive and can be useful for other researchers evaluating similar connectivity solutions from a sustainability perspective. Therefore, the consortium will follow an open data approach.

Additionally, a series of scientific publications in different conferences and journals, and a set of public deliverables will be generated along the project.

The data and documents produced during the COMMECT project are expected to be useful for research communities, decision-makers, rural communities and main actors in the agro-forest and transport industries. It could also be the case that the COMMECT project reuses data from previous projects. However, it cannot be specified which data will

be reused and what is its origin due to the early stage of the project. This will be further explained in *version 2* of the DMP (D7.4, due at M18).

3. Data Management

3.1. Internal Data Management

It is important that proper data format and storage plans are defined throughout and after project. Due to the early stage of the project, specific data formats or expected size of the data for all collected data cannot be fully described. Therefore, we provide here a preliminary plan for storage for each of the living labs, which will be updated and completed in deliverable D7.4 (*Data Management Plan - Version 2*), with due date in M18 (February 2024).

Figure 1 presents a scheme summarizing how the consortium plans to handle and store the data. The consortium distinguishes between two different groups:

- *Data collectors*: partners gathering the **raw data** from the end-users, in their own labs' facilities, and in the living labs. The data in this case will be internally stored at each of the partners' databases. Collected data is only kept in the data collector databases and not shared with the rest of the partners, the data format in this stage is not defined, and it is left to the criteria of the data collecting partner.

- *Data processors:* partners using the **post-processed and anonymized data** stored in a shared database that should be agreed between the partners. Access to the data will be given only to those partners that require it. It is recommended that post-processed data is shared with the same format in which it will be made publicly available, i.e., following the guidelines from Section 3.2 to provide open access to research data. Partners sharing data within a LL will agree on which is the most convenient database to share the data. They will also provide an API for the DST to access the mentioned database in a later stage of the project. The storage plan of each LL will be further specified later in this section.

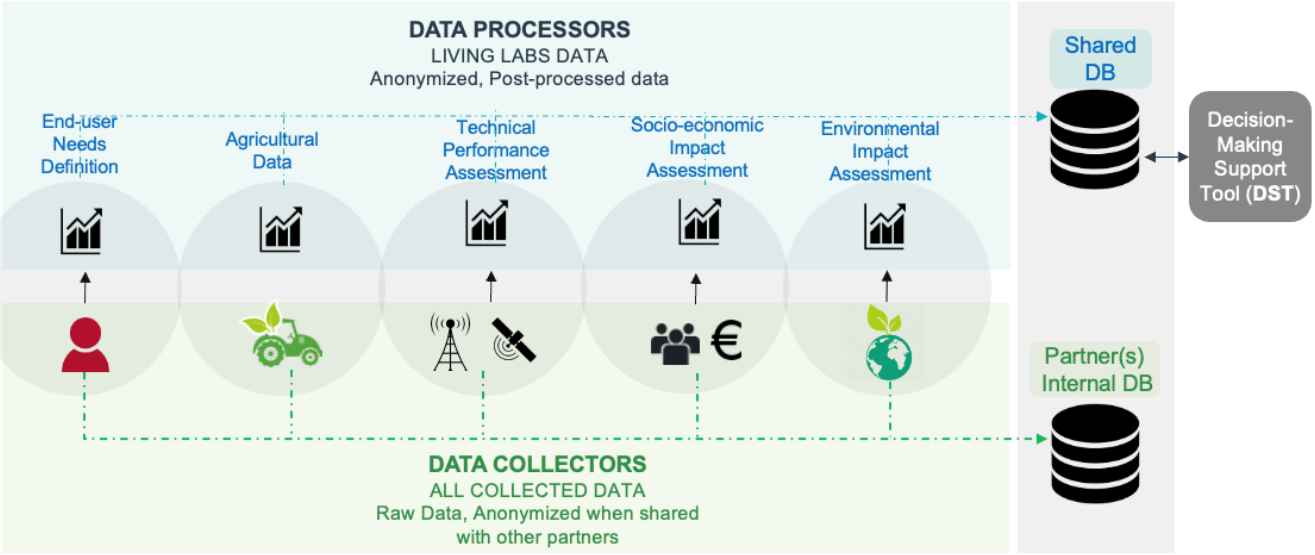


Figure 1. Data Flow Summary

During the first phase of the project, the focus will be in collecting end-user needs for the definition of use cases and requirements. This will be done through multiple interviews,

surveys, and workshops with the end users from the 5 different LLs. Data collectors in this case will be the responsible partners for each LL, while data processors will be the partners using the anonymized data to define the end-user requirements (from all technical, environmental, and socio-economic perspectives) and use cases. Data collectors will always be responsible of meeting the European GDPR and anonymizing data before sharing it with data processors. Before meeting with the end-users, data collectors should prepare the information sheet of the project, consent forms and privacy notice, i.e., provide end-users with terms and conditions in order to make to ensure that it is possible to use and share their inputs for any purpose such as promotion, distribution, exploitation or commercialization.

In the following stages of the project most of the collected data will be gathered in the different LLs for the technical performance, socio-economic impact, and environmental impact assessment. Only post-processed data is expected to be stored in the shared database.

In the following we detail what type of data is expected to be collected in each LL, and preliminary ideas on how it will be managed.

3.1.1. LL N.1 - Luxembourg – Digitalisation of Viticulture

In the living lab viticulture, we would like to support winegrowers in their daily work and make viticulture more resilient to climate change by the timely collection and provision of relevant data. We focus on data collection which, in a nutshell, is related to winegrowers' management practices, image data of vineyards and weather information. Especially for weather data the connectivity of the in-field sensors is the main challenge and will be solved within the project. Automated data transfer, data processing, severity calculations and potential alarms are defining the data format, flow and processing chain.

In the first phase of the project, the data is mainly collected by each partner separately and stored locally at their facilities. The Living Lab Luxembourg will focus on the user needs provided by a survey, which will be analysed in the first month of the project and updated continuously. Consequently, one data set is the anonymized answers about user needs, following GDPR, which are collected by IBLA.

At LIST, mainly satellite data (Sentinel-2) will be downloaded in the early stage and processed and in the following season UAV data will be collected. These data (only higher level products for satellite data; not raw data) are stored on the LIST fileserver, which provides access to all LIST employees which are involved in the project. Luxsense will also collect UAV data and terrestrial panorama images giving information on the detailed status of the vine canopy. Sensor data (temperature, leaf wetness) will be collected by LIST using several communication technologies (from terrestrial LoRaWAN to satellite backhauling).

3.1.2. LL N. 2 - Norway – Connected Forestry

The aim of the Forest LL is to support the forest industry with efficient and environmentally friendly connectivity-based solutions for their operations. The forest industry value chain contains major activities such as planting, care, thinning, logging, transport, and sawing. From a tree is planted to it is logged the time span is roughly 70-80 years. From initial discussions with key stakeholders, we learned that the logging activities are most interesting to focus on and improve. This activity lasts for roughly one to two months for the given geographical area. During this intensive period there are different aspects to include in the operation and data management such as: 1) Decision support for operation and maintenance, 2) Real time situation awareness and 3) Remote machine steering. For the former, this is relevant for more efficient logging

and for avoiding cutting rare biotopes. For the latter, remote steering (using AR/VR tools) by operators located away from the forest site mitigates reduced recruitment to the industry and risk of operator injuries.

The Forest LL partners suggest that the use of partner operation sensitive data is most convenient for internal data storage and only anonymized data useful for the development of the DMP is shared with the COMMECT platform. This is a requirement for collecting operation-sensitive data. However, some anonymized technical data for the decision support tool could be opened and shared. Telenor is the connectivity provider in the Forest LL and a transportable “on wheels” 5G private network solution is anticipated to apply for the trials related to logging operations. Environmental data regarding energy saving and other environmental indicators from digitalization and online-supported operations could be relevant to share. Some socio-economic impact data is also relevant to share such as social effects for municipalities related to deploying private networks for educational purposes or for the citizens and local businesses in general. Per now we do not expect to include user-generated data related to the operations. The Forest LL will return with specifications of the data formats later in the project.

3.1.3. LL N. 3 - Denmark – Connected Livestock Transport

The Living Lab in Denmark focuses on improving the connectivity experienced in the livestock transport sector. It will mainly test connectivity solutions suitable for different operations (loading/unloading of animals or over-the-route transportation), adopting several communication technologies (e.g., public cellular networks or satellite). During the project, the partners involved will investigate over-the-route 5G coverage assessment, multi-connectivity solutions to overcome coverage holes, or deployment of on-demand local green broadband network with terrestrial or satellite backhaul, dedicated for the provisioning of on-demand coverage and/or capacity during the loading/unloading of animals.

This LL will collect end-user data during the end-user needs definition phase, but this data will be mostly stored by AU, following GDPR when applicable. The rest of the partners will only receive an anonymized summary of the end-user needs. The LL will not collect specific information gathered by the trucks (sensor data – temperature, CO2 levels, etc. –, location), but will rather use the trucks to gather radio performance data along the main transport routes. Therefore, collected data will be mostly technical, socio-economic, and environmental data. Each partner will store collected data in their own servers. It is still unclear how much data, if any, will need to be shared between the different partners. Therefore, it is still being discussed which is the most convenient platform for processed data sharing in case it is needed.

3.1.4. LL N. 4 - Türkiye – Smart Olive Tree Farming

The Living Lab in Türkiye focuses on smart olive tree farming. The Living Lab Türkiye aims to support the determination of the most effective control time for diseases and pests by using early warning systems. The soil and other conditions in olive farms can be closely monitored using Turkcell’s XG networks that enable irrigation optimization, disease risk management and frost risk management, and optimum spraying decisions.

Collected data from Living Lab Türkiye covers end-user, technical and agriculture data. The end-user data is collected through workshops and surveys. All end-user data will be anonymized. Technical data comes from Turkcell XG’s RAN networks. Key Performance Indicators (KPIs) are use-case specific and confidential. The raw data

will be processed by the use case's requirements, and processed data will be shared with the analysis results. The agricultural data will be generated from the sensors of olive tree gardens. The sensors can be early warning system output, whether information or soil analysis. The Living Lab Türkiye's agricultural data API follows all the standards of the RESTful services; therefore, the implementation should be straightforward in any language.

3.1.5. LL N. 5 - Serbia – Sustainable Agriculture and Preservation of Natural Environment

The Living Lab in Serbia will focus on sustainable and digitized agriculture practice that will minimize negative effect on the environment and provide sustainable energy production for agriculture needs. Collected data during Living Lab deployment will be gathered from different sources like meteorological stations, soil moisture sensors, microphones and cameras. Proprietary service agroNET (https://digitalfarming.eu/) will be used for data visualization and data analysis trough the different expert modules in order to help end users in decision making.

All No. 5 LL data will be stored in DNET database. Storing data in a centralized database allows easy access and management of the information. By making the data available through an API, partners can easily retrieve and use the data in their own systems, without the need for manual data transfer. The data format used will be in accordance with "Smart Data Models"[1], a global program that establishes a standard for data formatting and management. This ensures that the data is organized and structured in a way that is easily understood and usable by all partners. To ensure the security of the data, access to the data for project purposes will be restricted to partners with appropriate credentials. This ensures that only authorized parties have access to the data and prevents unauthorized access or misuse. Overall, this approach allows for efficient data management and sharing among project partners, while also maintaining the security of the data.

3.2. Open Access to Research Data

Outputs formally released by the consortium and not compromising any legitimate commercial or other exploitation interest, will be shared according to the open science requirements described in the Horizon Europe Programme Guide, for making research data Findable, Accessible, Interoperable, and Reusable (FAIR) [3]. In the COMMECT project this mostly includes anonymized end-user data, technical data gathered in the experimental campaigns, as well as data regarding the socio-economic and sustainability aspects. Additionally, most of the project deliverables and scientific publications will be public. Data will be made publicly available through the Zenodo platform[2], a well-known repository under the European OpenAIRE program and operated by CERN, which allows researchers to upload and make openly available research datasets, publications, software and reports.

[1] https://smartdatamodels.org/

[2] https://zenodo.org/

3.2.1. Findability

To make the COMMECT data consistently findable and citable, all collected data that can be made publicly available and all the public documents of the project, will be **identified by a Digital Object Identifier (DOI).** DOIs are persistent identifiers that will provide long-term findability of any digital object on the Internet. Even though a Uniform Resource Locator (URL) is a unique reference to a digital object, it is not persistent, and if a digital object is removed from a certain website or location on the Internet, the URL will no longer be valid. To overcome this, persistent identifiers will consistently point to the object regardless its location. The URL of a certain object will be stored in the corresponding DOI and updated whenever the object's location is changed. For the COMMECT data:

- *Experimental datasets:* To make the collected data publicly available, this will be uploaded to Zenodo, which will generate a DOI for the corresponding dataset.
- *Public deliverables:* Public deliverables will be also uploaded to Zenodo, which will generate the corresponding DOI.
- *Scientific publications:* Those publications accepted by academic publishers will be assigned with a DOI, an International Standard Serial Number (ISSN), or a Published Item Identifier (PII), depending on the publisher itself and the editors' open access strategy (green or gold). Whenever possible, COMMECT will follow the green open access strategy.

Links to the public deliverables, scientific publications and datasets uploaded to Zenodo with the corresponding persistent identifier, will also be made available in the COMMECT project website[3].

Metadata

Rich metadata will be associated to all scientific publications, deliverables and processed data that it is planned to be made publicly available. Descriptive metadata containing information about the content of the data will be included, especially for scientific publications, deliverables, and end-user data. For datasets, structural metadata including information on how the data is arranged to facilitate a smooth understanding and use of it. All metadata will have associated keywords that suit the specific type of data. This will facilitate the process of making research data open access.

For collected and generated data (technical performance, socio-economic and environmental assessment), metadata will depend on the discipline, and will be included in the description field of metadata, as it is specified below. The COMMECT project considers five types of collected data:

- **Anonymized End-user**
- **Agricultural**
- **Technical**
- **Socio-economical**
- **Environmental**

As for the specific format of metadata, Zenodo is compliant with the DataCite Metadata Schema 4.4 [4], which is a de facto standard for describing datasets and the project will therefore follow the (metadata) standard(s), for the specific repository.

[3] https://www.horizoneurope-commect.eu/

The COMMECT partners are requested to fill out the following fields when uploading data to Zenodo:

- For scientific publications and project deliverables:
  - **DOI:** If the original published provided one already
  - **Publication Type**
  - **Publication Date**
  - **Title:** title of the publication
  - **Author(s):** name(s) of the author(s) involved in the development of the document
  - **Description:** abstract of the publication
  - **Language**
  - **Journal/Conference:** partners should select the appropriate section from the Zenodo system and fill in the information of the publisher.
  - **Persistent identifier(s)** will be created automatically and attached to the metadata file created by Zenodo.
- For datasets:
  - **Publication date:** date for which the dataset is being made publicly available.
  - **Title:** descriptive title for dataset
  - **Authors:** authors/company involved in the data collection/processing
  - **Description:** data format, data content, number of columns, reference to other datasets/similar datasets used for inspiration, and other information considered relevant for the understanding of the data.
  - **Keywords**

3.2.2. Accessibility

As previously mentioned in Section 3.2, the COMMECT project will use Zenodo to make data publicly available. Additionally, by providing DOIs for all uploaded data, Zenodo guarantees a lifetime of at least 20 years, and has a best-effort policy for migration of content in the case of closing operations [5]. This means that all COMMECT data will be kept in a trusted repository of open access research data that will be fully and permanently accessible.

However, as explained in Section 2, some of the data collected may be kept confidential. Which data will remain confidential is still to be decided and will be further clarified in a later stage of the project when there is a clear definition of the data to be collected. Additionally, the DST tool is expected to be exploited after the COMMECT project lifetime, and therefore, making it completely open, goes against the legitimate interests of some of the partners. The early stage of the project makes it difficult to define whether this tool will or not be exploited and which partners will benefit from it. For the same reason it is not possible to specify if an embargo would be applied to give time to publish or seek protection of intellectual property.

In any other case, for the data uploaded to Zenodo, it is planned to be anonymized, and it is not expected to have restrictions on use. Every person will be able to access it through the Zenodo website. Therefore, a data access committee will not be needed in this case.

As mentioned above, the decision-making support tool (DST) that the COMMECT project expects to provide as an outcome may be exploited after the project. In that case, code will not be open source, and third parties may have to pay a license to use the tool. If this situation occurs, it will have to be discussed whether the members of the consortium get a free license to use the tool, for how long does the license hold, and other practicalities. This can only be discussed in a later stage of the project. Other

software being outcome of the project activities (e.g., simulators, implemented algorithms, processing programs, etc.) may be made publicly available with open-source code. This will also be further specified in a later version of the DMP.

3.2.3. Interoperability

Due to the early stage of the project, and since data collection has just started (gathering end-user needs) and formats are not fully clear yet, it is difficult to state the level of interoperability of the data collected during the project. The project partners can confirm that, as above mentioned, metadata will be used following the standards of the different types of collected data so that data can be easily exchanged and reused within and across disciplines. Where possible, the documentation of the data will include descriptions with references to other datasets. The metadata in Zenodo will point to other relevant datasets, publications, etc. The consortium is evaluating the different options for data format as explained above and will include the final decision in *version 2* of the DMP. However, we anticipate that all openly accessible data will be uploaded in a commonly accessible format and encoding. It is not expected that the COMMECT project will use uncommon or project-specific ontologies or vocabularies.

At the time of defining data formats, it will also be considered whether data from previous projects will be reused. In that case, interoperability will be increased by either adapting the data formats from the previous projects or by including qualified references[4].

3.2.4. Reusability

The data gathered along the project is also expected to be useful for other EU projects within the HE framework, as well as for researchers addressing the digitalization of the agro-forest and transport industries. To facilitate data usage and understanding of the different datasets, tools and code made publicly available, the COMMECT partners will follow the guidelines:

- **Datasets:** all dataset that will be made publicly available will follow the guidelines explained in the previous sections, including metadata and DOIs in every case. This will allow future users to understand what data is contained in the datasets, in which format it is, when and where was it collected, etc. If there are any public articles using the data, the link to these will also be included in the metadata.
- **Scientific publications:** whenever possible, partners will try to publish with open access strategy, making information available to any other user that may be interested in reproducing or reusing the work, and validating the results obtained during the COMMECT project.
- **Code/tools:** for all code/tools uploaded to the Zenodo AFAIK repository, there will be a readme file with information on methodology (how to install, how to use, etc.), variables definition, etc.

Data will be uploaded to Zenodo, which is a well-known open access platform, so it will be available for third parties for at least 10 years after the project has ended [5]. Data will also be licensed using standard reuse licenses according to the obligations set out in the Grant Agreement. This rule is kept generic and flexible so that it Is possible to adapt it to the real needs during the project implementation.

[4] https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/

4. Data Security

4.1. Internal Data Security

For internal storage, partners collecting and storing the data are responsible of keeping version control and storing data in a safe location. Section 3.1 explained the storage plans that will be followed by the different LLs, ensuring a secure storage place.

4.2. External Data Security

For publicly available data, as it has been previously explained in Section 3.2, data will be stored in Zenodo. According to Zenodo, in the unlikely case that they would have to close operations, they have a best-effort policy for migrating all content to other suitable repositories. Thanks to the existing DOIs, that Zenodo itself provides, all uploaded data will not be affected. Zenodo also has a best-effort policy for data protection and security, and they explained the different methods they use for that in [6]. Their efforts focus on protecting their users accounts and the integrity of the datasets uploaded to their repository. Additionally, COMMECT will designate a data responsible for all collected and generated data or documents. The responsible will be also foreseeing that all collected data is stored safely in Zenodo with the corresponding metadata and DOIs.

5. Other Research Outputs

5.1. Deliverables

Links to the public deliverables uploaded to Zenodo will be available in the project website. A total of 32 deliverables, out of the 41 that will be generated during the project, will be publicly available in both platforms. The plan to make these files available following the FAIR principles has been explained in Section 3.2. This is done with the aim of disseminating the project results and making them available to the scientific community and any other community interested in consulting, reproducing, or reusing them. A full list of the COMMECT deliverables can be found in *Deliverable 7.1 – Project Handbook for Quality Assurance and Risk Management*.

5.2. Scientific Publications

All publications of the COMMECT will be provided as Open Access (OA), either immediately with publishing or with an embargo of maximum 6 months. The publications will include all relevant research outputs, such as data, software, scientific knowledge, models, etc. For peer-reviewed publications, each R&D partner reserved an allocated budget for open access publications in high-ranked journals and conferences. Access to non-peer-reviewed publications will be open and immediate. In both cases, the publications and related metadata will be shared via Zenodo and cross-linked with the project public website. Further to this the addendum to the publication agreement, provided by the European Commission will be used.

Due to the early stage in which the project is, it is not possible to provide a list of scientific publications. However, several scientific publications are envisioned as part of the outcome of the COMMECT project.

6. Allocation of Resources

No immediate costs are expected from making the HE COMMECT project results openly available and meeting the FAIR principles. Zenodo allows safe and free of charge data storage for the future in CERN's Data Centre for as long as CERN exists. In the case of unforeseen expenses for making the COMMECT project data publicly available, according to the Grant Agreement, these are eligible for reimbursement.

7. Data Management Responsible

The responsible partner for data management in the COMMECT project is Aalborg University (AAU). The role of AAU will mostly consist in ensuring that the FAIR principles are followed by the different partners and that the publicly available data follows the agreements stated above in this document or in later versions of it.

7.1. Data Responsible

A data responsible will be pointed out for every set of collected or produced data/deliverable/publication along the project. Each data responsible will mostly supervise that data management guidelines are followed. The responsible will supervise the internal storage of data and will be in charge of making sure that everything is anonymized, if necessary, that all partners involved in the corresponding activity are keeping the agreed format and sharing it in the selected open access platform (Zenodo). The responsible of each data set is specified in Table 1 of Appendix 1, which will be updated along the project.

8. Conclusion

This deliverable presents the first version of the Data Management Plan in the COMMECT project. Collected data along the project is expected to include end-user, technical, socio-economic, and environmental types of data. For internal data storage and management, data will be handled per Living Lab, as per agreed by the partners involved in each of the Living Labs. The document also presents available following the Findable, Accessible, Interoperable and Reusable (FAIR) principles. Ethical aspects of the project have been discussed in a specific project deliverable (D8.1) and have therefore not been addressed in this document.

The DMP is a living document and will be further updated in D7.4 – Data Management Plan *version 2,* which is due in February 2024.

Appendix 1: Data Responsible

Table 1. Data Responsible for all types of generated data during the COMMECT project. This table will be updated during the project.

Data Type	Description	Responsible Partner	Contact Person[5]
Deliverables	Report on end-users' needs and relevant use cases	Responsible partner of the deliverable	Person leading the deliverable
Publication	Report on COMMECT requirements and KPIs	Main author	Main author
WP1 – LL Luxembourg	Data collected for user needs definition. (Interviews, surveys, workshops, etc.)	LIST	Miriam Machwitz miriam.machwitz@list.lu
WP1 – LL Norway	Data collected for user needs definition. (Interviews, surveys, workshops, etc.)	KI	Nis Valentin Nielsen nis.valentin.nielsen@klosser.no
WP1 – LL Denmark	Data collected for user needs definition. (Interviews, surveys, workshops, etc.)	AAU	Melisa López mll@es.aau.dk
WP1 – LL Türkiye	Data collected for user needs definition. (Interviews, surveys, workshops, etc.)	TOB	Muzaffer Kerem Savran muzafferkerem.savran@tarimorman.gov.tr
WP1 – LL Serbia	Data collected for user needs definition. (Interviews, surveys, workshops, etc.)	DNET	Srdjan Krco srdjan.krco@dunavnet.eu
Socio-economic Data	Data collected for socio-economic impact assessment. (Interviews, surveys, workshops, etc.)	TNOR	Per Jonny Nesse per-jonny.nesse@telenor.com
Environmental Data	Data collected for environmental impact assessment.	SeAMK	Iida Viholainen Iida.Viholainen@seamk.fi

[5] Data management responsible may change over the project implementation. For the moment, we have considered the LL leaders as main responsible of the data collected so far in each LL, unless another responsible was already identified.

References

[1] Data Guidelines to Open Policy and Open Research in Horizon Europe. Available Online. Last Visited: November 2022. https://open-research-europe.ec.europa.eu/for-authors/data-guidelines/

[2] Horizon Europe Data Management Plan Template. May 2021, Available Online. Last Visited: January 2023. Available Online : https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/reference-documents;programCode=HORIZON *(Data Management Plan)*

[3] Guidelines on FAIR Data Management in Horizon 2020. Available Online. Last Visited: January 2023. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[4] DataCite Metadata Schema. Information available online. Last Visited: February 2023. https://schema.datacite.org/

[5] Zenodo Policies. Information available online. Last Visited: February 2023. https://about.zenodo.org/policies/

[6] Zenodo Infrastructure – Security. Information available online. Last Visited: February 2023. https://about.zenodo.org/infrastructure/