



DATA MANAGEMENT PLAN

31 MARCH 2023

MANLIO BACCO (CNR), SABRINA BRIZIOLI (CNR), GIANLUCA BRUNORI (UNIFI), VALENTINA COLCELLI (CNR)
TIZIANA NADALUTTI (UNIFI), GINA PAVONE (CNR)



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them



D8.1 DATA MANAGEMENT PLAN

Project name	Maximising the CO-benefits of agricultural Digitalisation through conducive digital ECoSystems
Project acronym	CODECS
Horizon Europe Topic ID	HORIZON-CL6-2021-GOVERNANCE-01-22
Project ID	101060179
Website	https://www.horizoncodecs.eu/
Document Type	Deliverable
File Name	D8.1. Data management plan
Status	Final
Dissemination level	Public
Date of creation	31 March 2023
Keywords	Data Management, Ethics, Open Science, Gender perspective, Agriculture, FAIR data, RDM, DMP
Authors	Manlio Bacco (CNR), Sabrina Brizioli (CNR), Gianluca Brunori (UNIFI), Valentina Colcelli (CNR), Tiziana Nadalutti (UNIFI), Gina Pavone (CNR)
Work Package Leader	WP8 - Coordination
Project Coordinator	University of Pisa



**Co-funded by
the European Union**

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

VERSION 1.0

History of changes			
Version	Date	Changes	Authors(s)
0.1	21 February	First draft	Gina Pavone, Sabrina Brizioli, Valentina Colcelli
0.2	3 March	Internal Review, administrative information added	Tiziana Nadalutti, Manlio Bacco
0.3	20 March	Document review after consortium check and suggestions	Gina Pavone, Sabrina Brizioli, Tiziana Nadalutti
0.4	28 March	Types removed and formatting with EC template	Tiziana Nadalutti
1.0	31 March	First version release approved and released	Gianluca Brunori, Tiziana Nadalutti

Project Number: 101060179

Project Acronym: CODECS

Project title: Maximising the CO-benefits of agricultural Digitalisation through conducive digital ECoSystems

Project starting date: fixed date: 1 October 2022

Project end date: 30 September 2026

TABLE OF CONTENTS

1	Project factsheet	3
1.1	Project beneficiaries.....	3
1.2	CODECS DMP deliverable factsheet.....	5
1.3	Project Overview.....	5
2	Data Management Plan in CODECS project	6
2.1	Scope of this DMP.....	6
2.2	Overall structure of the DMP in CODECS project.....	6
3	Data Summary	7
3.1	Types of data.....	7
3.2	Work packages data overview.....	8
3.2.1	WP1 - Dissemination, Exploitation, Communication and Outreach (DECO) strategy.....	8
3.2.2	WP2 - Fostering capabilities: towards a conceptual and operational framework.....	10
3.2.3	WP3 - Analysis of digital ecosystems.....	11
3.2.4	WP4 - Costs and benefits analysis – assessment of the potential of digital technologies in agriculture.....	13
3.2.5	WP5 - Demonstration of Costs and Benefits of digitalisation and technology assessment.....	15
3.2.6	WP6 - Policy analysis and roadmap.....	17
3.2.7	WP7 - CODECS platform.....	19
3.2.8	WP8 - Coordination.....	21
4	Data Management inside the consortium	22
4.1	File naming.....	23
4.2	Folder structure.....	23
4.3	Selection Criteria for long-term preservation.....	24
5	FAIR Data Management	24
5.1	Findability.....	25
5.1.1	Enriching metadata.....	27
5.2	Accessibility.....	30
5.3	Interoperability.....	31
5.4	Reusability.....	31
6	Other research outputs	32
6.1	Software.....	32

6.2	Communication materials	33
7	Estimated costs	33
8	Data security	34
8.1	Roles and responsibilities	34
8.2	Data storage and back-up	34
8.3	Quality assurance processes	35
9	Legal and ethical aspects	35
9.1	Analysis of the information provided on personal data	36
10	Gender perspective	36
10.1	The rationale for the gender dimension within the DMP	36
10.2	Analysis of the information provided on the gender perspective	39
11	Revisions of the Data Management Plan	40

1 Project factsheet

Project name	Maximising the CO-benefits of agricultural Digitalisation through conducive digital ECoSystems
Project acronym	CODECS
Horizon Europe Topic ID	HORIZON-CL6-2021-GOVERNANCE-01-22
Project ID	101060179
Project website	https://www.horizoncodecs.eu/
Granting authority	European Research Executive Agency
Project starting date	1 October 2022
Project end date	30 September 2026
Project duration	48 months
Project website	https://www.horizoncodecs.eu/
Document type	Deliverable
Title	CODECS Data Management Plan
Authors	Gina Pavone, Sabrina Brizioli, Valentina Colcelli
Date of creation	March 2023
Status of the document	Final version
Version number	1.0
Dissemination level	Public
Internal reviewers	Gianluca Brunori, Tiziana Nadalutti, Manlio Bacco
Keywords	Agriculture, FAIR data, RDM, DMP

1.1 Project beneficiaries

Beneficiary n. and role	Organisation	Acronym	Country
1 - COO	UNIFI	UNIFI	Italy
2 - BEN	ACTA	ACTA	France

Beneficiary n. and role	Organisation	Acronym	Country
2.1 - AE	INSTITUT FRANCAIS DE LA VIGNE ET DU VIN	IFV	France
2.2 - AE	INSTITUT DE L'ELEVAGE	IDELE	France
2.3 - AE	IFIP-INSTITUT DU PORC ASSOCIATION	IFIP	France
3 - BEN	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR	Italy
4 - BEN	SZECHENYI ISTVAN EGYETEM	SZE	Hungary
5 - BEN	ASSOCIATION EUROPEENNE POUR L'INFORMATION SUR LE DEVELOPPEMENT LOCAL	AEIDL	Belgium
6 - BEN	CENTRO INTERNAZIONALE DI ALTISTUDI AGRONOMICI MEDITERRANEI	CIHEAM-IAMB	Italy
7 - BEN	EIGEN VERMOGEN VAN HET INSTITUUT VOOR LANDBOUW- EN VISSERIJONDERZOEK	EV ILVO	Belgium
8 - BEN	ELLINIKOS GEORGIKOS ORGANISMOS - DIMITRA	ELGO-DEMETER	Greece
9 - BEN	CESKA ZEMEDEL'SKA UNIVERZITA V PRAZE	CZU	Czechia
10 - BEN	GEOPONIKO PANEPISTIMION ATHINON	AUA	Greece
11 - BEN	BIOSENSE INSTITUTE - RESEARCH AND DEVELOPMENT INSTITUTE FOR INFORMATION TECHNOLOGIES IN BIOSYSTEMS	BIOS	Serbia
12 - BEN	UNIVERZA V LJUBLJANI	UL	Slovenia
13 - BEN	ZEMNIEKU SAEIMA	ZSA	Latvia
14 - BEN	UNIVERSITAET HOHENHEIM	UNI HOHENHEIM	Germany
15 - BEN	NEW EDU NO	NEWEDU	Slowakia
16 - BEN	INSTITUT NATIONAL DE RECHERCHE POUR L'AGRICULTURE, L'ALIMENTATION ET L'ENVIRONNEMENT	INRAE	France
16.1 - AE	INSTITUT NATIONAL D'ENSEIGNEMENT SUPERIEUR POUR L'AGRICULTURE, L'ALIMENTATION ET L'ENVIRONNEMENT	Institut Agro	France
17 - BEN	SUSTAINABLE IRELAND CO-OPERATIVE SOCIETY LTD	SUST	Ireland
18 - BEN	NODIBINAJUMS BALTIC STUDIES CENTRE	BSC	Latvia
19 - BEN	AG FUTURA TECHNOLOGII DOOEL SKOPJE	AGFT	Macedonia

Beneficiary n. and role	Organisation	Acronym	Country
20 - BEN	CONSORZIO TUTELA PECORINO TOSCANO	Pec.Toscana	Italy
21 - BEN	EESTI MAAULIKOOL	EMU	Estonia
22 - BEN	CSITA	CSITA	Czechia
23 - BEN	ISO-TECH SP Z OO	ISO-TECH	Poland
24 - BEN	UNIVERSIDAD DE CORDOBA	UCO	Spain
25 - BEN	UNIVERSIDAD DE ALMERIA	UAL	Spain
26 - BEN	ASOCIACION DE ORGANIZACIONES DE PRODUCTORES DE FRUTAS Y HORTALIZAS DE ALMERIA	COEXPHAL	Spain
27 - BEN	STICHTING WAGENINGEN RESEARCH	WR	The Netherlands
28 - BEN	Laboratoire d'Innovation Territorial Ouest Territoires d'Elevage	LIT OUSTEREL	France
29 - AP	THE JAMES HUTTON INSTITUTE	Hutton	UK

1.2 CODECS DMP deliverable factsheet

Task number: 8.3

Task title: Ethics, Open Science, data management and gender perspective

Due date (month): M6

Following Updates: D8.2 (M18); D8.3 (M36); D8.4 (M48)

Responsible partner: UNIPI (with support of CNR)

Work Package: WP8 Coordination

Dissemination level: PUBLIC

Contributors: all partners

1.3 Project Overview

The CODECS project aims at improving the collective capacity to understand, assess and foresee the full range of benefits and costs of farm digitalisation. It will co-develop, together with farmers and Agricultural Knowledge and Innovation System (AKIS) actors, user-friendly approaches, methods, and tools able to document the co-benefits and the costs of technologies applied to real contexts. CODECS project will develop a vision of “sustainable digitalisation” which will

contribute to a multi-level transition that links social, economic, and ecological aspects together, by adopting an innovative system based and actor-centred approach, and an action research methodology based on the coordination of 21 Living Labs through interdisciplinary and transdisciplinary teams. The research will study the role of 'digital ecosystems' in the rate of uptake of digital technologies and in the distribution of their costs and benefits; will develop and test dedicated indicators to monitor the level of digitalisation and foresee its potential costs and benefits. In addition, the project is intended to test and demonstrate digital technologies in 21 Living Labs through physical and virtual demonstration and will work on policy analysis and the production of specific policy tools.

The activities will be implemented through 8 work packages: Three work packages (WP3, WP4, and WP5) will be dedicated respectively to analysis of Digital Ecosystems [WP3], analysis of Costs and Benefits [WP4], Demonstration and Assessment activities [WP5]. These three work packages, that organise the core activities of Living Labs, will be supported and complemented by five horizontal work packages: Dissemination, Exploitation, Communication, and Outreach (DECO) strategy [WP1], Conceptual framework and methodology [WP2], Policy tools [WP6], Digital support tools [WP7], and Overall coordination [WP8].

2 Data Management Plan in CODECS project

2.1 Scope of this DMP

The Data Management Plan is a document that outlines the management and handling of research data both on a daily and long-term basis, establishing early-stage choices about the main aspects of the lifecycle of research outputs. In the CODECS project, the DMP is intended to plan and make careful choices about the data that will be used to achieve the project's goals, with the aim of ensuring that data is managed securely and in accordance with the highest standards of quality and ethics, in compliance with all regulations involved.

In the light of the general framework outlined by the European Commission, which promotes open and collaborative science, the data management of this project adheres to the principles of Open Science. In accordance with Horizon Europe guidelines¹, this DMP establishes the rules for responsible management of research data in line with the FAIR principles of 'Findability', 'Accessibility', 'Interoperability' and 'Reusability', and according to the principle of "as open as possible, as closed as necessary", in compliance (also) with the European General Data Regulation Policy (GDPR).

2.2 Overall structure of the DMP in CODECS project

This document is intended to be a single point of reference for all data-related issues and in general for all CODECS research outputs, from day-to-day data management issues (e.g. the organisation of information and data safety strategies) to all aspects of data sharing and dissemination of results.

Information on the data that the consortium partners will handle during the project have been collected through a dedicated questionnaire, which was answered by WP leaders. The answers provided were used to both structure the present Data Management Plan and the Personal data policy.

¹ See Horizon Europe Programme Guide: V2.0 – 11.04.2022

https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf

The CODECS DMP is established following the Horizon Europe DMP template², and thus containing the following sections:

- Data summary, in which all relevant information on data types and their collection or gathering process are described for each WP.
- The data management in CODECS project, where information on the organisation of data management among consortium partners are provided.
- FAIR data management, with information on each component of the FAIR acronym, included details on provisions for metadata, access to data and protection of personal data, licensing, and the modalities through which Open Access to publications will be carried out.
- Other research output, where a strategy will be outlined for managing and sharing the software and communication materials produced within the project.
- Estimated costs, where the allocation of the necessary resources will be planned.
- Data security, where choices to secure all project information and the quality assurance process to be put in place are outlined.
- Legal and ethical aspects, in which details on the management of personal data are given.
- Gender perspective, where specific actions are planned on the basis of indicators developed within the consortium.
- Revisions of the DMP, where roles and responsibilities will be specified.

3 Data Summary

3.1 Types of data

In CODECS project data will be classified according to the source used, whether primary or secondary, and according to their content, whether texts, tables, or multimedia. Information will be included on provenance or data collection processes and on all the aspects relevant for data reuse and how to facilitate it. Besides, a further distinction will be made between personal and non personal research data in order to detail specific treatment related to personal data. A personal data policy will be shared among the CODECS project partners for the lawful processing of personal data.

In order to answer the research questions and carry out the activities foreseen in CODECS project, both primary and secondary data will be used, in a combination of qualitative and quantitative analysis.

In CODECS the following types of data will be described:

- Documents (reports, briefs, deliverables, presentations)
- Datasets
- Multimedia
- Software

Based on the type of collection, a distinction will be made between:

² See Data Management Template (HE): V1.0 – 05.05.2021, downloadable from:

<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/reference-documents;programCode=HORIZON>

- Surveys
- Interviews and focus groups
- Measurements (e.g. in-field data collection) and data inventory

With Measurement data we refer to in-field measurement, administrative records, field notebooks, farm management information software (FMIS)

- Experiments (e.g. activities in Demo Farms and Living Labs, as planned in WPs 3, 4 and 5)
- Literature

3.2 Work packages data overview

Below are the characteristics of the data to be handled by each WP.

3.2.1 WP1 - Dissemination, Exploitation, Communication and Outreach (DECO) strategy

Lead beneficiary: AIEDL

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- measurements
- experiments
- literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- Numeric (e.g., tables, counts, measurements)

- ✓ Audiovisual (e.g., images, sound recordings, video, presentations)
- Instrument-specific (e.g., equipment outputs)
- Software

Primary data collection mode

Interviews will be realised with project partners, Living Lab members, demo farms, Knowledge Accelerator actors, etc. Such interviews will be recorded and used to produce communication materials. Besides data will be collected to measure the DECO impact.

Surveys will be conducted among project partners, Living Lab members, demo farms, KA actors, etc. with the purpose of understanding optimal ways to reach CODECS target audiences.

Source of secondary data

Materials and resources (e.g. images, templates for slides or infographics) may be taken from the following databases: Artgrid, Adobe, Freepik, Pexels, Pixabay, YouTube Library for Creators, project partners, AIEDL’s internal archives, Canva.

Purpose of the data

In this WP the purpose of collecting **primary** and **secondary** data is to develop the dissemination and communication (DECO) strategy, as well as to produce communication and dissemination materials (videos, interviews, briefings, practice abstracts).

File Formats

- | | | |
|---|-----------------------|-----|
| ✓ | DOC/DOCX | CVS |
| ✓ | JPEG/PNG/SVG | TSV |
| ✓ | PDF | ODT |
| ✓ | PPT/PPTX | TXT |
| ✓ | MPx | XLS |
| ✓ | MP4/MOV/WMV/WEBM..../ | |

Expected size

More than 2 MB, considering both primary and secondary data.

Software needed

Microsoft Office suite; Adobe; Sendinblue (for mailing lists), Canva (for photo editing and media graphics elaborations)

3.2.2 WP2 - Fostering capabilities: towards a conceptual and operational framework

Lead beneficiary: WUR

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- measurements
- experimental
- ✓ literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- ✓ Numeric (e.g., tables, counts, measurements)
- ✓ Audiovisual (e.g., images, sound recordings, video, presentations)
- Instrument-specific (e.g., equipment outputs)
- ✓ Software

Primary data collection mode

In order to carry out the WP2 activities, primary data will be collected through face-to-face interviews, an email survey, workshop setting. The data resulting from the collection will be organised in a Microsoft SQL server.

Source of secondary data

Scientific literature, policy documents, national statistical offices, international statistical offices (OECD, FAO), other institutions and/or research facilities.

Purpose of the data

In this WP the collection of both **primary** and **secondary** data aims at designing a conceptual framework of what sustainable development means in different agricultural contexts, as well as to develop a methodology to provide guidance for the digital transition in these contexts. In addition, the collected data will contribute to reporting about the living lab setup and results will provide insights into synergies and differences across the living labs.

File Formats

✓ DOC/DOCX	✓ CVS
JPEG/PNG/SVG	TSV
✓ PDF	ODT
PPT/PPTX	TXT
MP3x	✓ XLS
✓ RDF	✓ OWL

Expected size

More than 2MB.

Software needed

Microsoft Office suite; Adobe; GIS software; Reference software (Endnote, Mendeley); SurveyMonkey; VARI; Microsoft SQL server.

3.2.3 WP3 - Analysis of digital ecosystems

Lead beneficiary: UAL

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- ✓ measurements
- experiments
- literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- ✓ Numeric (e.g., tables, counts, measurements)
- ✓ Audiovisual (e.g., images, sound recordings, video)
- ✓ Instrument-specific (e.g., equipment outputs)
- ✓ Software

Primary data collection mode

The WP3 activities will be carried out using primary data collected through interviews, focus groups, surveys and measurements.

Source of secondary data

Secondary data will derive from scientific literature, government statistics, reports from other research facilities.

Purpose of the data

In WP3 primary and secondary data will be collected for the purposes of mapping the socio-ecological context. Measures, criteria and indicators, graphical representations will be created for Living Labs. Components of the socio-technical process modelling will be represented. Besides, based on the data of the different tasks to be performed, a comparative analysis of digital ecosystems will be carried out. Also, data on the role of AKIS systems will be gathered.

Secondary data is also needed to realise a comparative assessment of digital ecosystems.

File Formats

✓	DOC/DOCX	CVS
✓	JPEG/PNG/SVG	TSV
✓	PDF	ODT
✓	PPT/PPTX	TXT
	MPx	XLS

Expected size

More than 2MB.

Software needed

Microsoft Office suite; OpenOffice; Adobe.

3.2.4 WP4 - Costs and benefits analysis – assessment of the potential of digital technologies in agriculture

Lead beneficiary: INRAE

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- ✓ measurements
- ✓ experiments
- ✓ literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- ✓ Numeric (e.g., tables, counts, measurements)
- ✓ Audiovisual (e.g., images, sound recordings, video, presentations)
- ✓ Instrument-specific (e.g., equipment outputs)
- ✓ Software

Primary data collection mode

Primary data will be collected through measurements and inventories (for ex. Life Cycle inventories), web surveys, in depth interviews, focus groups and video testimony.

Source of secondary data

Scientific literature, reports, existing inventory databases and statistical indicators will contribute to the Cost & Benefits (C&B) analysis of the use of Digital Technologies at farm level, within the CODECS Living Labs.

These data will contribute to the Cost & Benefit analysis (eco, environmental, social) of the use of digital technologies at farm level, within the CODECS Living Labs.

Purpose of the data

Data collection will allow performing analysis on:

- Cost & Benefit perceptions.
- Environmental impact analysis of the use of digital technologies in the Living Lab's focal action situation based on comparative life cycle assessment.

- Economic performance of the digitalisation process, based on budgeting/partial budgeting methods and Net Present Value analysis (at farm level).
- The costs and benefits for the farmer/user in terms of social aspects, including an assessment of the social impacts of digital agriculture.

File Formats

✓	DOC/DOCX	CVS
✓	JPEG/PNG/SVG	TSV
✓	PDF	ODT
✓	PPT/PPTX	TXT
✓	MPx	✓ XLS
✓	.MOV	

Expected size

More than 200MB (both primary and secondary).

Software needed

Microsoft Office suite, NVIVO, SPSS, SIMAPro.

3.2.5 WP5 - Demonstration of Costs and Benefits of digitalisation and technology assessment

Lead beneficiary: ACTA

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews

- ✓ surveys
- ✓ measurements
- ✓ experiments
- ✓ literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- ✓ Numeric (e.g., tables, counts, measurements)
- ✓ Audiovisual (e.g., images, sound recordings, video, presentations)
- ✓ Instrument-specific (e.g., equipment outputs)
- ✓ Software

Primary data collection mode

Primary data will be collected through FMIS (Farm Management Information Systems), DSS (Decision Support Systems), agricultural machinery, online surveys, online/video/phone interviews with different actors, such as agtech providers, agricultural engineers, researchers.

Source of secondary data

The sources of secondary data to be used for WP 5 activities are:

- national statistical offices;
- scientific literature;
- agricultural technical institutes (applied research) literature;
- engineering consulting firms;
- private companies.

Purpose of the data

Primary and secondary data will be collected in order to create a demonstration network and to produce demonstration guidelines and set-up protocols. Data that will be collected in WP5 will depend on WP8 inputs about LLs characterization.

File Formats

- | | |
|----------------|-------|
| ✓ DOC/DOCX | ✓ CVS |
| ✓ JPEG/PNG/SVG | TSV |
| ✓ PDF | ODT |
| ✓ PPT/PPTX | TXT |
| MPx | ✓ XLS |
| ✓ .SPH | |

Expected size

More than 2 MB, considering both primary and secondary data.

Software needed

Microsoft Office suite; Adobe; GIS software.

3.2.6 WP6 - Policy analysis and roadmap

Lead beneficiary: BSC

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- measurements
- experiments

literature

Content of the data

✓ Text (e.g., field or laboratory notes, survey responses)

Numeric (e.g., tables, counts, measurements)

Audiovisual (e.g., images, sound recordings, video)

Instrument-specific (e.g., equipment outputs)

Software

Primary data collection mode

Original data will be collected through surveys, interviews, questionnaires, and self-administered questionnaires, semi-structured interviews, workshop discussions.

Source of secondary data

Scientific literature, reports, policy recommendations, and publications on the state of digitalisation (and rural digitalisation) in partner countries.

Purpose of the data

Through the review of existing policies regarding digitalisation in agriculture and rural areas, a report in Policy environments for digitalisation in Europe will be produced.

A survey to AKIS actors will study how they can support the diffusion of evidence about digital technologies Cost & Benefit.

File Formats

✓	DOC/DOCX	CVS
	JPEG/PNG/SVG	TSV
✓	PDF	ODT
	PPT/PPTX	TXT

MP3/M4a/WAV

✓ XLS

Expected size

Around 200MB for both primary and secondary.

Software needed

Microsoft Office suite.

3.2.7 WP7 - CODECS platform

Lead beneficiary: AUA

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- measurements
- experiments
- literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- ✓ Numeric (e.g., tables, counts, measurements)
- ✓ Audiovisual (e.g., images, sound recordings, video, presentations)

✓ Instrument-specific (e.g., equipment outputs)

✓ Software

Primary data collection mode

Surveys and interviews will be realised to collect the CODECS platform user needs.

Source of secondary data

Secondary data will be collected from already existing inventories of other projects delivering information about digital farming technologies.

Purpose of the data

The WP7 will produce the CODECS platform, intended as a single-point-of-access for:

- a meta-inventory of digital technologies
- the Digital technology impact assessment toolkit
- the Sustainability Performance demonstration toolkit

The CODECS platform will act as a quality-checked database where all the project outputs and results will be gathered and showcased. A data entry tool will be built within WP7 activities.

File Formats

- | | |
|----------------|-------|
| ✓ DOC/DOCX | ✓ CVS |
| ✓ JPEG/PNG/SVG | TSV |
| ✓ PDF | ODT |
| PPT/PPTX | TXT |
| MPx | ✓ XLS |
| ✓ JSON | ✓ PNG |
| MPG | ✓ WMV |
| ✓ MP4 | |

Expected size

More than 100MB.

Software needed

Microsoft Office suite; Adobe; Visual Studio Code, Mongo DB as DBM software.

3.2.8 WP8 - Coordination

Lead beneficiary: UNIPI

Origin of the data

- ✓ primary data
- ✓ secondary data

Type of collection

- ✓ interviews
- ✓ surveys
- measurements
- experimental
- literature

Content of the data

- ✓ Text (e.g., field or laboratory notes, survey responses)
- ✓ Numeric (e.g., tables, counts, measurements)
- ✓ Audiovisual (e.g., images, sound recordings, video, presentations)
- Instrument-specific (e.g., equipment outputs)
- Software

Primary data collection mode

The original data to be produced by this WP will be:

- textual data, produced through reports, project deliverables, timelines and work plan preparation
- numeric and tabular data, derived from the update on financial aspects and project progresses

Source of secondary data

Grant Agreement and consortium management.

Purpose of the data

The data will be used to ensure:

- efficient management, financial monitoring and reporting;
- compliance with research ethics and open science, and a proactive approach to gender dimension;
- the coordination of Living Labs and monitor the learning process.

File Formats

✓ DOC/DOCX	✓ CVS
✓ JPEG/PNG/SVG	TSV
✓ PDF	ODT
✓ PPT/PPTX	TXT
MPx	✓ XLS

Expected size

More than 100MB.

Software needed

Open Office, Adobe, Microsoft Office Suite.

4 Data Management inside the consortium

Given the multidisciplinary nature of the project and the international dimension of the consortium, adopting a clear organisation of information is essential for the proper management of research data. To this end, a clear convention for

file naming has been adopted, as well as an agreed-upon folder structure, documented with a read-me file located at the root folder.

4.1 File naming

All files will be findable using the following file naming convention:

DocumentType_Title_VersionNumber(if needed)_Dateofrelease.extention

For example: PRES_DMPinCodecs_14122022.ppt

The date format is DDMMYYYY

The document type are shortened as follows:

- DATA - short name DATA
- DELIVERABLE – DEL
- EXPERIMENT - short name EXP
- INTERVIEW - short name INT
- LITERATURE - short name LIT
- PROJECT MANAGEMENT TOOL - short name PMT
- MULTIMEDIA - short name MM
- POLICY BRIEF - short name PB
- PRESENTATION - short name PRES
- REPORT - short name REP
- SURVEY - short name SURV
- WORKING PAPERS - WPP

4.2 Folder structure

In the folder structure, a division by WP will be followed and work in progress will be separated from the stable versions. The following structure will be followed:

- CODECS
 - General
 - 0_Bibliography
 - 1_Proposal&Evaluation
 - 2_Grant&Consortium_Agreements
 - ConsortiumAgreement
 - GrantAgreement
 - 3_Delivered
 - FinalDeliverables (for Deliverables accepted by the Commission)
 - Reports (for accepted version of Reports to the Commission)
 - WP number
 - Activities or content, for example:
 - Interviews
 - Country

- Year
- Surveys
- Country
 - Year
 - Literature
 - Training
 - ...
 - Deliverable D.X.Y (for ongoing work)

4.3 Selection Criteria for long-term preservation

In order to facilitate the reproducibility of the results and to increase the reusability of the data, these must be preserved in the long term through a reliable repository. For the choice of repository and its characteristics, see the section FAIR data management.

As CODECS project will use different types of data, the following rules will apply for long-term preservation:

- Primary data collected in the project will be preserved.
- Secondary data, already available from other sources and provided with a persistent and unique identifier (PID), will not be preserved.
- Secondary data that, for the purposes of the project, are rearranged into original datasets (as for instance statistical data, even if data themselves have not been changed), will be preserved.
- Datasets whose data will be cited by any of the publications, reports, or deliverables, will be preserved.

To facilitate versioning, at different stages of the research life-cycle, depending on the level of processing, data in CODECS project may be referred to as:

Raw data: unprocessed data, eventually with full personal information (non-anonymized data).

Processed data: recodes, transformations, selections, or enrichment of the anonymized data, which are all captured and described in metadata and documentation files.

Stable data: processed data, on which no new data are added. These are preferably selected for long-term preservation.

In addition, personal data may be referred to as:

Pseudonymised (personal) data: raw data in which any personal identifiers are removed, but whose link to the data allowing identification is located separately.

Anonymised (personal) data: raw data in which any personal identifiers are removed.

5 FAIR Data Management

The acronym FAIR - standing for Findable, Accessible, Interoperable, Reusable - and the FAIR principles are widely acknowledged in the scientific community as the model to follow to optimise the dissemination and exploitation of scientific resources. The aim of FAIR principles is to increase the possibility of finding, understanding, and using research data for anyone who can benefit from it.

This section describes the choices and actions for the CODECS project to be in line with the FAIR principles.

In accordance with the Grant Agreement, research data within CODECS project will follow the principle of “as open as possible, as closed as necessary”.

5.1 Findability

A dedicated community in the Open Access **repository** Zenodo will be set up in order to facilitate the discoverability of CODECS research outputs and to guarantee long-term preservation.

The policy of the Zenodo repository is available at the following link: <https://about.zenodo.org/policies/>

A guide for advanced search is provided here: <https://help.zenodo.org/guides/search/>

For the unique identification of products, on Zenodo a **Digital Object Identifier (DOI)** can be issued for each deposited product. Alternatively, an already registered DOI can be entered (for example if the publisher already provided a DOI for a publication). In addition, the use of **ORCID** as researchers’ PID is supported.

Persistent identifiers also enable the proper citation of research artifacts and a standard citation is provided by Zenodo in order to facilitate the reuse and repurpose of the dataset while ensuring proper credit and recognition to the author(s).

Zenodo's **metadata** is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments. Supplementary metadata will also be added and attached in .txt files whenever necessary to accurately describe the resource and make it fully understandable and reusable. Whenever possible the Data Documentation Initiative (DDI) metadata standards³ will be used to document survey data and other data emerging from the activities of the Living Labs (social, behavioural, economic). For agricultural concepts, terms, definitions, the AGROVOC⁴ multilingual controlled vocabulary will be considered.

In Zenodo the metadata of each record is **indexed** and searchable directly in the repository search engine immediately after publishing; the metadata of each record is also sent to DataCite servers during DOI registration and indexed there. Moreover, Zenodo's records are indexed in OpenAIRE, the Open Access Infrastructure for Research in Europe helping researchers report their publications to the EC Participant Portal and comply with the European Commission Open Access Policy. In fact after the records’ metadata are harvested from OpenAIRE, they also become available on the EC participant portal.

The Zenodo community administrator will also take care that all the research outputs are properly linked to the CODECS project in OpenAIRE, so as to facilitate monitoring of results.

In addition, and in particular when required by partners' institutional policies, products may also be deposited in institutional repositories.

Below the list of Zenodo's metadata, with description.

Macroarea	Metadata field	Description
	Upload type*	Can be: publication, poster, presentation, dataset, image, video/audio, software, lesson, physical object, workflow, other

³ <https://ddialliance.org/>

⁴ <https://agrovoc.fao.org/browse/agrovoc/en/>

Macroarea	Metadata field	Description
Basic information	Digital Object Identifier*	Can be cited an already existing one or newly issued
	Publication date*	Format: YYYY-MM-DD
	Title*	The title of the resource
	Author(s)*	The name(s) can be complemented with affiliation and ORCID
	Description*	Textual explanation and details
	Version	Mostly relevant for software and dataset uploads. Any string will be accepted, but semantically-versioned tag is recommended
	Language	Primary language of the record
	Keywords	Selected words to facilitate discoverability
Licence	Access right*	Can be: open access; embargoed access; restricted access; closed access
	Licence*	The licence specifying the condition for the reuse of the resource
Funding	Grants	Allows to specify the Research Funding Organization and grant number
Related/alternate identifiers	Related identifiers	Allows to specify identifiers of related publications and datasets
Contributors	Contributors	The name(s) can be complemented with affiliation and ORCID
References	Reference	Supports linking other resources (also external to Zenodo)
Journal	Journal title	Can be specified: journal title; volume; issue; pages
Conference	Conference title	Can be specified: conference title; acronym, dates, place, website, session, part
Book/Report/Chapter	Publisher	Can be specified: publisher; place; ISBN; book title; pages
Thesis	Awarding university	Can be specified: awarding university, supervisors
	Supervisors	The name(s) can be complemented with affiliation and ORCID
Subjects	Subjects	Here subjects can be specified from a taxonomy or controlled vocabulary. Each term must be uniquely identified (e.g. a URL)

*Mandatory fields

5.1.1 Enriching metadata

Depending on the type of data managed within the CODECS project, the following additional metadata will be added.

Documents

All CODECS **documents** will display the following information on one of the initial pages:

Field	Content / description
Project name	Maximising the CO-benefits of agricultural Digitalisation through conducive digital ECoSystems
Project acronym	CODECS
Horizon Europe Topic ID	HORIZON-CL6-2021-GOVERNANCE-01-22
Project ID	101060179
Project website	https://www.horizoncodecs.eu/
Document type	Report / Deliverable / Briefing / etc. (see the complete list of document types at paragraph 4.1)
Title	The title of the document
Status	Draft / Final version /Final Version accepted by the Commission
Version number	
Dissemination level	Internal/Public...
Date of creation	DDMMYYYY
Date of release	DDMMYYYY
Author(s)	
Contributors	
Internal reviewers	
Work package Leader	

Field	Content / description
Project Coordinator	
Keywords	

Dataset and multimedia

To supplement the information expressed through the Zenodo repository metadata, additional metadata will be attached with a documentation file in .txt format. In this metadata, all the information needed to understand and enhance the reusability of the dataset will be added. If metadata standards suitable for the resource, especially disciplinary ones, will be identified, this will be indicated in the documentation .txt file.

Additional metadata may include:

Metadata field	Description
Mode of collection	The procedure, technique, or mode of inquiry used to attain the data. Vocabulary: DDI controlled vocabulary on mode of data collection: https://ddialliance.org/Specification/DDI-CV/ModeOfCollection_3.0.html
Country/ies of reference	To which country(ies) does the resource pertain
Year(s) of reference	To which year(s) the content of the resource refers
Data collection start date	Indicates the start of data collection
Data collection end date	Indicates the end of the data collection
Last Update	Indicates the date the resource was last modified
Data type	Dataset / Multimedia
Data formats	E.g.text / numbers / images / 3D models / audio files / video files / / surveys / maps / scientific articles .txt; .csv; .shp; .pdf; .stadat; .wav; .mp3; mp4 etc.
Data origin	Primary or Secondary
Source	The source of secondary data reused in CODECS project. Ideally a PID or URL can be added to the original source.
Licence	The original licence of secondary data

Metadata field	Description
Software	If specific software is needed to access/process the data. If multiple are available, one is sufficient. Open Source or Free ones are preferred.
Data size	Known data size
Data version	Can be: raw data; pseudo anonymised personal data; anonymized personal data; processed data (see Personal data policy and paragraph 4.3)
Specific metadata for Video	<p>This may include:</p> <p>Type of product (interview, video-news release, documentary, stockshots, clips, etc); Director (if any); Place of the event; Start date of shooting; End date of shooting; Start date of distribution (if any) End date of distribution (if any); Script or short list with names and functions of the personalities filmed and clear identification (e.g.: from left to right; 2nd from left; etc.); Links and other useful information/website where the product can be viewed or downloaded; Technical aspects (such as format, sound transmission, etc)</p>
Specific metadata for images	<p>This may include:</p> <p>Date when the photo has been taken; Place where the photo has been taken; Event description; Description of each photo with names and functions of the personalities photographed and clear identification (e.g.: from left to right; 2nd from left; etc.); Name of the photographer; Links and other useful information/website where the product can be viewed or downloaded; Technical aspects (such as format, resolution, etc.)</p>
Specific metadata for audio	<p>This may include:</p> <p>Date when the audio was recorded; Place where the photo has been taken; Event description; names and functions of the personalities audio-recorded; Name of the interviewer/speaker/presenter; Links and other useful information/website where the product can be heard or downloaded; Technical aspects (such as format, etc.)</p>

Software

The Software that will be produced within the project to collect or analyse data will be shared through Zenodo repository, within the CODECS project community. It will be appropriately documented according to current standards (see section "Other research outputs"), and versioning procedures will be applied, if needed.

5.2 Accessibility

For long-term preservation and to increase the discoverability of the project results, the catch-all repository Zenodo will be used. Thus the record metadata are always publicly available and licensed under **public domain** and no authorization is needed to retrieve them. At the same time, metadata contains the DOI of the record and they allow the user to reach the data itself.

Access to literature

CODECS project will comply with European Commission Open Access mandate on scientific literature by depositing a copy of each published article derived by the project activities that will undergo a peer-review process. All deposited articles will be assigned an open access right at the time of publication (no embargo period will be applied to scientific publications).

Green Open Access (i.e., self-archiving a version allowed for deposition in OA) will be the preferred route, although a specific budget to publish in Gold Open Access journals is reserved (only for full Open Access Journals, see paragraph "Estimated costs" for APCs).

Beneficiaries will ensure that they retain sufficient rights to provide Open Access to the scientific literature they will produce in the CODECS project.

The corresponding author of the paper published is responsible for depositing the open access version of the paper that is compliant with the publisher policy and the European Commission mandates.

As soon as possible after paper acceptance, and at the latest at the time of publication, the corresponding author should deposit in compliance with EC Open Access mandate and publisher Open Access policy, the version of the article that is permitted for open access in the Zenodo repository, include reference to the CODECS grant in the metadata, and set the open access to the file containing the machine-readable version of the article.

The use of Open Research Europe publishing platform will be encouraged, as well as the beneficiaries' participation in open peer review processes.

All articles should contain the reference to the project, possibly in the "funding" or "acknowledgment" sections. Suggested form is the following:

"This [article/chapter/report/result/equipment] is part of the CODECS project - "Maximising the CO-benefits of agricultural Digitalisation through conducive digital ECoSystems". The project has received funding from the European Union's Horizon Research and Innovation Actions under the HORIZON-CL6-2021-GOVERNANCE-01 call, grant agreement No 101060179."

Specific training will be provided on all issues related to Open Science and Research Data Management (see the section on training in the paragraph "Quality assurance processes").

Access to data

The data collected or produced within the CODECS project will be made accessible through the Zenodo repository (see the paragraph 5.1 "Findability").

In accordance with the Grant Agreement, unrestricted and royalty-free access will allow third parties and the general public mining, exploitation, reproduction, and dissemination of research data generated by CODECS, thus maximising the impact on the scientific community.

Should the need arise, embargo periods will be clearly declared, accounting the need to balance openness and exploitation, protection of scientific information, commercialization, IP Rights, security, and preservation questions.

Open access will be granted to all data selected for deposition as soon as possible and considering proper versioning, or at the latest as soon as they are considered stable. See paragraph 4.3 "Selection criteria for long-term preservation" for further details.

As the CODECS project aims to collect primary data through interviews and surveys, all personal data will be treated in compliance with ethical standards and in full respect with the provisions of the GDPR.

In all cases where personal data will be handled for research purposes, these will only be shared following anonymisation or pseudonymisation process, as described in the article 4(5) of GDPR.

If the data contain sensitive or person-identifying information, appropriate choices will be made (anonymization using specific software such as AMNESIA - <https://amnesia.openaire.eu/> - or restricted or closed access).

If the restricted access option is chosen, guidance will be provided in the deposited item, clarifying who can request access, under which conditions, and to whom to address the request.

5.3 Interoperability

In order to enrich metadata and to achieve interoperability with other data, the resources listed in initiatives such as the RDA Metadata Standards Catalog (<https://rdamsc.bath.ac.uk>) and the FAIRsharing registry of standards and semantic artifacts (<https://fairsharing.org/search?fairsharingRegistry=Standard>) will be taken into account for the retrieval of specific controlled vocabularies.

In addition to disciplinary related resources, Zenodo uses vocabularies for licences (provided by Open Definition), funders (by FundRef), and grants (by OpenAIRE).

To enrich the contextual knowledge about the data, while depositing a new research output, whenever suitable an attempt will be made to link and reference other related products with the appropriate metadata field, with the goal to create as many meaningful links as possible between (meta)data resources.

5.4 Reusability

Beneficiaries will commit to attach additional documentation whenever it is necessary to add information to the metadata provided, in order to make the resources more understandable and thus to increase opportunities of data reuse.

Documentation

Within the consortium, the documentation will be stored together with the data, in the corresponding data subfolder. For deposition and public sharing, the documentation will be attached in the form of a readme file in .txt format. The readme files could eventually complement with information on methodology, codebooks, machinery settings, data cleaning, analyses, variable definitions, units of measurement, etc.

The project outputs will be reusable by third parties also after the end of the project. To this end, data and metadata are retained for the lifetime of the Zenodo repository.

The chosen repository also allows DOI versioning allowing to:

- edit/update the record's files after they have been published.
- cite a specific version of a record.

- cite all of the versions of a record.

Licences

To permit the widest reuse possible, CODECS outputs will be made freely available with a licence of the Creative Commons framework.

The deposited data will be made available under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights.

Immediate open access is provided to the deposited publication via the Zenodo repository, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a licence with equivalent rights; only for monographs and other long-text formats, the licence may exclude commercial uses and derivative works (e.g. CC BY-NC, CC BY-ND).

If any of the data to be collected will be subject to commercial exploitation, the DMP will be immediately updated with full details of the type of data involved and the type of exploitation that will be applied.

Standard citation

When the data will be deposited and access will be granted for reuse, a standard citation will be provided in order to facilitate reuse and repurpose of the dataset while ensuring proper credit and recognition to the author. Zenodo provides standard citations in different styles and which can be exported in different formats, all containing basic metadata (author, title, date) and a resolvable DOI.

6 Other research outputs

No physical products such as samples, new materials, reagents or similars are expected in the CODECS project. This section will therefore give details on the management of other digital products resulting from the project work and activities, in addition to the data, scientific literature and the so-called grey literature (i.e. everything beyond academic publishing in the narrow sense).

6.1 Software

All relevant software and computer code that will be produced in CODECS project to collect and analyse data will be properly documented and openly shared through the Zenodo repository.

Licences are essential to establish the conditions for re-use, modification or distribution of the software as well. Once it has been ascertained who owns the intellectual property of the software created (the researcher(s) or the institution they work for), it will preferably be released under a Free and Open Source Software (FOSS) licence⁵ between the ones listed by the Open Source Initiative: <https://opensource.org/licenses/>

⁵ Morin A, Urban J, Sliz P (2012) A Quick Guide to Software Licensing for the Scientist-Programmer. PLoS Comput Biol 8(7): e1002598. <https://doi.org/10.1371/journal.pcbi.1002598>

For anyone to be able to read and understand the research software, the next points will be followed as far as possible:

- variables and functions names will be as self-descriptive as possible.
- proper code commenting will be realised while developing.
- a readme file in .txt format will be attached, including all irrelevant information (e.g. how to install and configure the software, if a full documentation is available and where it is possible to find it, if a quickstart guide has been released etc.).
- proper version control both of the software and of its documentation will be applied using Zenodo versioning functionalities.

6.2 Communication materials

One of the project goals is to enable a two-way communication between researchers, all the involved stakeholders, the widest scientific community and society at large. A variety of outputs are envisaged for dissemination purposes: articles, newsletters, reports and videos. Besides, social media accounts will be managed. For the protection and management of personal data, please refer to the ethics section. Whereas for the selection of material to be eventually deposited in Zenodo, please refer to the specified criteria (see paragraph 4.3).

Moreover, a project website will be developed for the project results to be publicised. The website server is hosted by Siteground with IP 35.214.213.135. The appropriate cookie policy and terms of use will be displayed on the website. Whenever possible, the DOIs of products published elsewhere and deposited on Zenodo will be referred. The project data, properly anonymized if needed, may also be displayed on the project website, always referencing the deposited version in the repository with the appropriate PID, and to facilitate findability and indexing, the use of descriptive metadata structured in the RDF-a or Microdata format will be considered.

7 Estimated costs

The following costs are estimated:

1. planned budget for long-term preservation: no costs are expected for using the Zenodo repository.
2. purchase of the following supplies and software for the communication and dissemination strategy:
 - Artgrid: 18,60 EUR/month
 - Freepik 8,06 EUR/month
 - Sendinblue 49 EUR/month
 - Total: 75,66 EUR/month, that is 3.631,68 EUR for the whole project (48M)
3. Publication costs: Gold Open Access will be considered when publishing in leading international journals. An estimated budget of 20.000 euros has been allocated for the APCs in fully Open Access publication venues.
 - Beneficiaries are aware that publication fees are only eligible when publishing in full open access journals.
 - The Green Open Access will be practised in case a subscription-based venue will be chosen for publication.
 - The Open Research Europe publishing platform will also be considered.
4. Other infrastructure or software costs that are not detailed here are considered in-kind by the consortium institutions.

8 Data security

This section addresses data storage during the project, the foreseen backup strategy and specifies who does what. Moreover, the initiatives and processes to ensure data quality are outlined.

8.1 Roles and responsibilities

Researchers, technicians and Living Lab participants will be in charge of **collection** of research data.

Both Living Lab coordinators, acting as supervisors, and researchers will deal with data quality assurance. With specific reference to personal data it is clear that the security of personal data and the compliance with the GDPR are in charge to the data controller (art.4, n.7 controller 'means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law').

If needed, institutional staff, such as data management staff, could be involved in the management of research data.

All data will be on servers that are password protected with secure online access and authentication measures.

The project coordinator (UNIFI) will be responsible for granting access to the data in the shared project folders, while the list of authorised staff will be provided by the project leader of each partner institution.

The project coordinator is responsible for the quality of the DMP and for its implementation during the project's lifetime.

Within the project consortium, all the involved beneficiaries are responsible for

- uploading all data related to the project into the designated data repository; verifying that sufficient metadata is provided;
- ensuring that data structure, naming, etc. comply with the DMP;

Besides:

- CNR-Ifac is responsible for managing personal data protection and ethical issues.
- UNIFI (with CNR-ISTI support) is responsible for the update of this DMP whenever any significant change occurs and at scheduled mandatory updates.

8.2 Data storage and back-up

The data will be stored in institutional servers and cloud services during the collection phases. All the data relevant for the activities of the consortium will be stored and managed by the Microsoft teams cloud service provided by the project coordinator (UNIFI).

This cloud environment is based on Microsoft Azure technology. The disaster recovery solution is based on the backup solutions offered by Microsoft Azure cloud services, which are GDPR compliant. In addition, the project coordinator will ensure at least monthly back-ups of all project data stored in non-MS infrastructure.

8.3 Quality assurance processes

Since CODECS project includes data collection from different stakeholders through living labs, specific training materials and activities are planned in order to ensure the highest level of quality. Training activities will be carried out periodically, with a focus on how to collect data and how to assure data quality.

In addition, manual control, **preprocessing and cleaning** steps will be implemented on unstructured raw data whenever necessary.

During the CODECS project proper training will be provided on **legal and ethical aspects and open science**. In particular, the beneficiaries will be trained on the different pre- and post- peer-review versions (preprint, Author's Accepted Manuscript - AAM - and the Version of Record - VoR), on the Rights Retention Strategy, and on the use of tools to check editorial policies such as Sherpa Romeo, specifying that publication fees are only eligible when publishing in full open access journals.

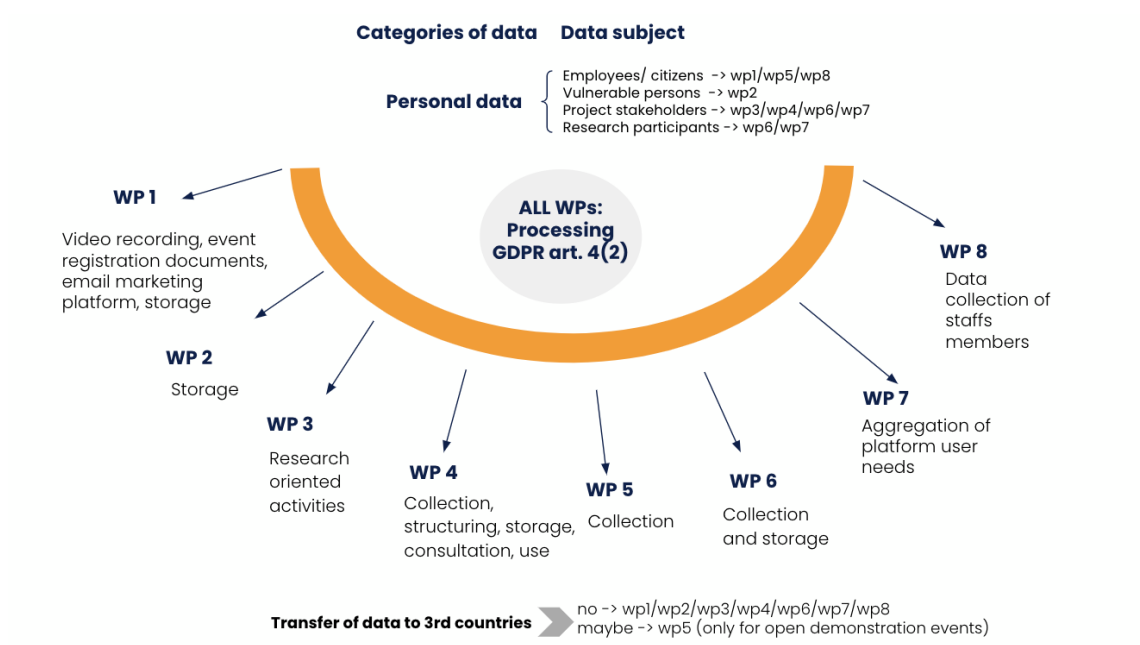
9 Legal and ethical aspects

An overview of the CODECS' partners approach to the processing of personal data is possible thanks to a preliminary survey aiming at estimating the data flow across the WPs of the CODECS project.

The figure above summarises the answers to the second part of the questionnaire addressed to the partners and dealing with the processing of personal data. It is worth pointing out the way the requirements and safeguards as indicated by the GDPR are set up by the CODECS partners in order to prevent and avoid personal data breach.

The figure resumes the kind of personal data, which will be collected, and provides an overview on the data flow to address the data protection in WPs, mainly dealing with: informed consent procedures and criteria to be safely involved in research activities, lawful and purposeful processing of personal data, estimation of subjects involved (i.e. data subjects, recipients, data controller, data processor) and secure storing.

Figure 1. Data flow for each WP



9.1 Analysis of the information provided on personal data

In the framework of CODECS activities, all the Consortium members collect personal data through projects. Personal data may be collected because they are the object of the project research activity, but also when they may be needed for purposes inherent to the project, but different from the research: for instance, the collection of lists of stakeholders' names, addresses for the project dissemination phase, etc. Most partners have stated their intention not to use special categories of data, except for two partners who are uncertain about their use.

Personal data will therefore be collected and protected in compliance with European and national legislation. This means that scientific research with personal data, especially related to people's health, will be guaranteed by an activity inspired by the integrity and high ethical standards.

All partners are aware of the necessity to obtain informed consent to treat personal data even though the procedures/criteria are various and must be properly addressed by the CODECS consortium. Partners demonstrate commitment by implementing targeted consent forms, information sheets and storing tools but common rules and strategies should be established within the CODECS consortium.

Considering the categories of data, some partners have expressly indicated the use of personal data, while others have simply expressed the likelihood of handling this category of data. When asked to specify the category of data to be used, partners effectively demonstrated to refer to personal data.

Starting from the analysis summarised in the figure above, which summarises the type of personal data that will be processed within the project, partners will receive a Privacy Policy to support them in processing the personal data collected, preceded by a training seminar for its correct implementation. Please note, just for the sake of the description of this document, that consent to the research activity is separate consent from consent to the processing of personal data for the research activity. They are two consents that invoke two disclosures for though partially overlapping purposes, and indeed consent for the research activity often contains and invokes the lawfulness and security in the processing of personal data for those who participate in the activity. For the above-mentioned reason, the Ethic Guide will be realised and shared among the partners with a specific focus on how to collect interviews. Further details on ethics are tackled in the Ethics guidelines.

10 Gender perspective

10.1 The rationale for the gender dimension within the DMP

The expert groups on legal and ethical issues within the CODECS project will work to monitor in a systematic and strategic manner the gender perspective in CODECS research and training activities. For this reason, information and data about gender are enclosed in the overview of data provided by the Data Management Plan. It is worth underlining that the questions referred to the gender composition/balance in the WP structure and do not concern the overall presence of women in CODECS.

Thus, some questions of the preliminary survey that were addressed to the CODECS partners dealt with the gender dimension in order to scrutinise the gender balance in the organising and working structure of WPs as well as to estimate whether gender participation occurs on equal footing.

The following table illustrates women's engagement in WPs and it attempts to investigate CODECS partners' attitudes to cast/ or not considerable attention on gender in research.

WP	Of the total of workforce, how many are women?	What roles do women perform?	Do you think gender diversity might improve performances?	Are there any strategies or policies to promote gender diversity?	Do you apply non-discriminatory and anti-harassment policy?	What are (if any) the main challenges to involving women?	Are you aware of any legal, institutional and political drivers or bottlenecks to women's involvement?
WP 1	Out of AEIDL's total workforce, 66% are women if looking at AEIDL's team working on CODECS, at the moment 80% are women	Team allocated to CODECS: For CODECS: project manager (at the moment); communication officers, policy officer	Yes	AEIDL has a Gender Equality Task Force and has conducted a recent audit of deployment and pay	Yes, there is an in-house staff experience that includes provisions for discrimination and harassment		No that we are aware of
WP 2	3 of 6	Lead of WP2, expertise on data, young professional who conduct literature reviews	Yes	No, but women are included based on their specific expertise	There is a system to report and get help at WUR with these kinds of occurrences	None	We are aware of the bottlenecks and take action if need to be
WP 3	Of UAL about 50%; of project team, about 50%	WP lead, IP in UAL, research	Not sure in this case, but in general, yes	Yes	Yes	They are not taken as seriously, mansplaining, assumed to speaking longer when they are not, etc.	In what? the project? or society at large. If the latter, how long of a list would you like?

WP	Of the total of workforce, how many are women?	What roles do women perform?	Do you think gender diversity might improve performances?	Are there any strategies or policies to promote gender diversity?	Do you apply non-discriminatory and anti-harassment policy?	What are (if any) the main challenges to involving women?	Are you aware of any legal, institutional and political drivers or bottlenecks to women's involvement?
WP 4	10 (without counting the LL coordinators/participants)	All types of roles	Yes	Gender equality plan in some institutions	Yes	N/A	No
WP 5	Roughly: 31 women over 83 people involved in WP5 task workforce	WP leader, task leader	Yes	Hiring more women and balance men/women salaries	Yes	More balanced society, more diversity in sharing ideas and during physical activities that need female point of view	No
WP 6	N/A	N/A	N/A	Gender equality plans	Yes	N/A	No
WP 7	25%	Active involvement in WP 7 tasks. (research and development)	Yes	Any strategies described in our organisations gender	Yes of course	No such challenges faced or anticipated in the future	No

WP	Of the total of workforce, how many are women?	What roles do women perform?	Do you think gender diversity might improve performances?	Are there any strategies or policies to promote gender diversity?	Do you apply non-discriminatory and anti-harassment policy?	What are (if any) the main challenges to involving women?	Are you aware of any legal, institutional and political drivers or bottlenecks to women's involvement?
		t related activities)		equality plan			
WP 8	60%	Women involved in WP8 work at project management	Definitely	No particular strategies	We intend to apply such a policy but for the moment each partner adopts its own strategy. Gender Equality Plans are in force for almost all the partners	A discussion on this topic should be open in the partnerships. As a matter of fact, anyway, a significant part of each staff is composed by women	No

10.2 Analysis of the information provided on the gender perspective

Almost all partners (6/8), with the exception of one of them, indicated the involvement of women in the activities of the WP even though the percentage of engagement varies from one to another.

Almost all partners (6/8), with the exception of one of them, clarified that women are involved in many/all kinds of roles in WPs: WP lead, professionals involved in research studies and multi-tasking activities.

The majority of partners (6/8) believes that gender diversity might improve performances, but just few of them (3/8) expressly indicated policies/strategies relating to it, that is gender equality plans or targeted audits. One partner suggested hiring more women and ensuring more balanced men/ women salaries. In this last case, it is not clear whether this answer expressed a planned measure or an overall achieving target. One partner answered this question negatively.

The majority of partners (6/8) is aware of the enactment of non-discriminatory and anti-harassment policies, but few partners (2/8) made reference to them.

It appears, partners did not estimate any challenge in involving women. Two partners underlined very general obstacles/stereotypes and suggested taking women into consideration.

The majority of partners (5/8) negatively answered the question about institutional and political drivers or bottlenecks to women's involvement, one partner showed awareness about them and another one seemed to referred to a series of drivers/impediments for women in society at large.

11 Revisions of the Data Management Plan

The present Data Management Plan will be mandatorily updated at M18 (March 2024), M36 (September 2025) ad M48 (September 2026) of the CODECS project. It will also be updated whenever any relevant change occurs.