



## Pathways towards a fair, inclusive and innovative Data Economy for Sustainable Food Systems

### **D7.3: Data Management Plan (DMP) - M36**

Lead Author: Mireille van Hilten (WR)  
Contributors: Chris Avdellas (RFF),  
Maria Gallego Rebellon (WR)



Funded by  
the European Union

## Document Information

Grant Agreement No.	101059473
Project Acronym	Data4Food2030
Project Title	Pathways towards a fair, inclusive and innovative Data Economy for Sustainable Food Systems
Type of action	RIA - Research & Innovation Action
Call ID & Topic	HORIZON-CL6-2021-GOVERNANCE-01-18: Mapping and improving the data economy for food systems
Start – ending date	1 September 2022 – 31 August 2026   48 months
Project Website	<a href="https://data4food2030.eu/">https://data4food2030.eu/</a>
Work Package	WP7: Project Coordination
WP Leader	Wageningen Research (WR)
Deliverable type <sup>1</sup>   Dissemination level <sup>2</sup>	Report   PU
Due Date	31 August 2025
Deliverable Submission Date	15 August 2025
Version	Final
Internal Reviewer(s)	Maria Gallego Rebellon (WR), Chris Avdellas (RFF)

### Disclaimer

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

### Copyright message

This document contains unpublished original work unless clearly stated otherwise. Previously published material and the work of others has been acknowledged by appropriate citation or quotation, or both. Reproduction is authorised provided the source is acknowledged.

<sup>1</sup> Please check DoA for the deliverable type. Available options are: R: Document, report; DEM: Demonstrator, pilot, prototype, plan designs; DEC: Websites, patents filing, press & media actions, videos, etc.; DATA: Data sets, microdata, etc; DMP: Data management plan; ETHICS: Deliverables related to ethics issues; SECURITY: Deliverables related to security issues; OTHER: Software, technical diagram, algorithms, models, etc.

<sup>2</sup> Please check DoA for the dissemination level. Available options are: PU – Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page); SEN – Sensitive, limited under the conditions of the Grant Agreement; Classified R-UE/EU-R – EU RESTRICTED under the Commission Decision No2015/444; Classified C-UE/EU-C - EU CONFIDENTIAL under the Commission Decision No2015/444; Classified S-UE/EU-S – EU SECRET under the Commission Decision No2015/444

## Data4Food2030 Consortium

	Participant organisation name	Short name	Country
1	STICHTING WAGENINGEN RESEARCH	WR	NL
2	REFRAME FOOD ASTIKI MI KERDOSKOPIKI ETAIRIA	RFF	EL
3	INSTITUT NATIONAL DE RECHERCHE POUR L'AGRICULTURE, L'ALIMENTATION ET L'ENVIRONNEMENT	INRAE	FR
4	EIGEN VERMOGEN VAN HET INSTITUUT VOOR LANDBOUW- EN VISSERIJONDERZOEK	EV ILVO	BE
5	INSTITUT FÜR ANGEWANDTE SYSTEMTECHNIK BREMEN GMBH	ATB	BE
6	WAGENINGEN UNIVERSITY	WU	NL
7	THE LISBON COUNCIL FOR ECONOMIC COMPETITIVENESS ASBL	LC	BE
8	VRIJE UNIVERSITEIT BRUSSEL	VUB	BE
9	NETCOMPANY-INTRASOFT SA	INTRA	LU
10	STICHTING EFFOST	EFFoST	NL
11	EUROPEAN FOOD INFORMATION RESSOURCE AISBL	EUROFIR	BE
12	NODIBINAJUMS BALTIC STUDIES CENTRE	BSC	PL
13	SZKOLA GLOWNA GOSPODARSTWA WIEJSKIEGO	SGGW	DE
14	FONDATION JEAN JACQUES LAFFONT, TOULOUSE SCIENCES ECONOMIQUES	TSE	FR
15	EURO POOL SYSTEM INTERNATIONAL (DEUTSCHLAND) GMBH	EPS	BE
16	ANYSOLUTION SL	ANYSOL	ES
17	ITC - INOVACIJSKO TEHNOLOSKI GROZD MURSKA SOBOTA	ITC	SI
18	WIRELESSINFO	WRLS	CZ
19	1001 LAKES OY	LAKES	FI
20	AGDATAHUB	ADH	FR
21	IFIP-INSTITUT DU PORC ASSOCIATION	IFIP	FR
22	UNIVERZA V MARIBORU	UM	SI
23	UHLAVA OPS	UHLAVA	CZ
24	LUONNONVARAKESKUS	LUKE	FI

## Glossary of terms and abbreviations

List of Abbreviations and Acronyms	
4TU	4 Technical Universities (Dutch collaboration)
CA	Consortium Agreement
CC0	Creative Commons Zero (license)
CC BY	Creative Commons BY - Credit must be given to the creator (license)
CERIF	Common European Research Information Format
CERN	"Conseil Européen pour la Recherche Nucléaire" - European Council for Nuclear Research
CESSDA	Consortium of European Social Science Data Archives
DANS	Data Archiving and Networked Services
DDI	Data Documentation Initiative
DE4FS	Data Economy for Food Systems
DOI	Document Object Identifier
DMP	Data Management Plan
DRYAD	Curated general-purpose repository that makes data discoverable, freely, reusable, and citable
DSI	Data Sharing Initiative is a collaborative framework where stakeholders, such as farmers, researchers, companies, and policymakers, share data using a common infrastructure, based on agreed ethical, legal, and technical principles
EML	Ecological Language
EO	Earth Observation
EOSC	European Open Science Cloud
ERP	Enterprise Resource Planning (system)
EU	European Union
FGDC-CSDGM	Federal Geographic Data Committee's (FGDC) Content Standard for Digital Spatial
FIWARE	Future Internet ware - open-source standards for Smart applications
FMS	Farm Management System
GA	General Agreement
GAIA-X	Project working on the development of a federation of data infrastructure and service providers for Europe ( <i>GAIA-X - Home</i> , n.d.)
Gb	Gigabyte
GDPR	General Data Protection Regulation
GGN	Global GAP Number
GIS	Geographical Information System
GLN	Global Location Number
GS1	Global Standard 1
IoF2020	The Internet of Food and Farm 2020 (EU project)
IPR	Intellectual Property Rights
ISO	International Standards Organisation
MS	Monitoring System, the 'Monitor' aiming to provide insights into the data economy in food systems, as part of the Data4Food2030 project
NIS	Security of Network and Information Systems – Directive; the first piece of EU-wide legislation on cybersecurity
WUR	Wageningen University and Research

## Table of Contents

<b>Document Information</b> .....	<b>2</b>
<b>List of Tables</b> .....	<b>6</b>
<b>Revision History</b> .....	<b>6</b>
<b>1 Introduction</b> .....	<b>7</b>
<b>2 Data Summary</b> .....	<b>8</b>
<b>2.1 Data categories</b> .....	<b>8</b>
2.1.1 Research data .....	8
2.1.2 Case data.....	9
2.1.3 Meta data .....	9
2.1.4 External data .....	9
2.1.5 Personal data .....	9
<b>2.2 Data re-use</b> .....	<b>10</b>
2.2.1 Purpose.....	10
2.2.2 Data described .....	10
<b>2.3 Data management procedures</b> .....	<b>12</b>
<b>3 FAIR data</b> .....	<b>15</b>
<b>3.1 Making data findable</b> .....	<b>15</b>
<b>3.2 Making data accessible</b> .....	<b>16</b>
<b>3.3 Making data interoperable</b> .....	<b>17</b>
<b>3.4 Increase data re-use</b> .....	<b>17</b>
<b>3.5 Other research outputs</b> .....	<b>18</b>
<b>4 Allocation of resources</b> .....	<b>19</b>
<b>6 Ethics</b> .....	<b>21</b>
<b>6.1 Informed Consent Procedure</b> .....	<b>21</b>
<b>6.2 Stakeholder groups</b> .....	<b>22</b>
<b>6.3 Gender-inclusive approach and integration</b> .....	<b>22</b>
<b>6.4 Data consent retraction procedure</b> .....	<b>22</b>
<b>7 Next steps</b> .....	<b>23</b>
<b>8 References</b> .....	<b>24</b>
<b>Appendix A: Data4Food2030 Information letter V1.1</b> .....	<b>25</b>

**Appendix B: Informed consent procedure .....28**

**Appendix C: F.A.I.R. Data management guidelines.....29**

**Appendix D: GDPR Compliance & Re-identification Risk Checklist.....30**

**Appendix E: Data management Actions per Data4Food2030 Work Package .....32**

## List of Figures

Figure 1 Input data categories and examples .....8

Figure 2 Research data publishing to Zenodo (preliminary) ..... 13

Figure 3 Data Management Plan development timeline.....23

## List of Tables

Table 1 Data collection characteristics per category ..... 11

Table 2 Resource allocation for data management..... 19

Table 3 Data management activities during project .....23

## Revision History

Document	Major revisions	Author
<b>DMP version 3 (M36)</b>	Updated entire document text for changes. Added FAIR data principles guidelines, Privacy checklist and data management actions list for consortium partners Removed data management sessions (item 3 on page 12) due to CEADS upcoming data sharing frameworks	Mireille van Hilten (WR), Chris Avdellas (RFF), Maria Gallego Rebellon (WR)
<b>DMP version 2 (M18)</b>	Added informed consent procedure Other updates and revisions	Mireille van Hilten (WR), Joep Tummers (WR), Maria Gallego Rebellon (WR)
<b>DMP version 1 (M4)</b>	First draft and first version	Mireille van Hilten (WR), Sjaak Wolfert (WR), Joep Tummers (WR), Maria Gallego Rebellon (WR), Pia Groenewolt (VUB), Daoud Urdu (WR), Dimitris Fotakidis (FSH)

## 1 Introduction

Several initiatives are undertaken to foster a Data Economy for Food Systems (DE4FS) from a technical perspective, such the development of a specific Agricultural Data Space, and a reference architecture of smart farm management systems (FIWARE). In the GAIA-X project representatives from politics, business and science in the EU create a proposal for the next generation of a data infrastructure for Europe as a secure, federated system that meets the highest standards of digital sovereignty while promoting innovation. These initiatives have technological focus, and actions by different stakeholders are needed to address social challenges as the DE4FS develops. A holistic assessment of the state-of-play, of data-driven innovation and data reuse of the drivers, barriers, and enablers of digital transformation, and of ongoing and expected trends, including behavioural ones is needed.

Data4Food2030 aims to improve the data economy for food systems by expanding its definition, mapping its development, performance, and impact to create new insights and opportunities. The project aims to enlarge the knowledge base with insights into the data economy for food systems and to develop a system that monitors the development, performance, and impact of the DE4FS on relevant EU policies. Drivers and barriers of the DE4FS are investigated by the project team and turned into opportunities and recommendations, in nine case studies and through stakeholder dialogues. Future scenarios and a roadmap will detail how to sustain the system and how to support policy development and accelerate the desired future state of the DE4FS.

The Monitoring System (MS), or the ‘Monitor’ observes the development, performance, and impact of the DE4FS to be used primarily by EU and Member State policy makers, but also other stakeholders. It is a platform connecting data sources to visualise the DE4FS by e.g., indicators dashboards, graphs, tables, and contain links to documents related to the knowledge base and a back-end system that handles all data interactions. Populated with data, the system provides the actual state-of-play of the DE4FS, and it is therefore particularly important that data in the MS are published in an open format. Several distinct data categories require data governance because they need to provide insights into the development of the DE4FS at micro, meso and macro level. Public and private data need to be managed differently, and the project may require handling personal data. An adequate Data Management Plan needs to be provided by the project coordinator in collaboration with the core project members.

“A Data Management Plan (DMP) is a formal document that provides a framework for how to handle the data material during and after the research project” (CESSDA Training Team, 2022) and it is a living document. As part of Task 7.4 the proper handling of data that is collected within the project is described in the DMP, first submitted in month 4 (M4) and will be continuously updated until the end of the project. The objective of this version of the DMP is to provide an update on how data from the project is managed (i.e., content, type, format, volume), which standards and methodologies can be used for data collection and management, how they can be exploited and how they can be made findable, accessible, interoperable and reusable (FAIR).

Previous EU data projects in the agri-food domain, such as IoF2020, Smart Agri Hubs (SAH) and AgROBOfood provide best practices for data management. However, data management repositories, such as Zenodo, have developed since and the team needs to revisit data management ways of working. The basis for this DMP is the Horizon template for a Data Management Plan (*Horizon Europe Data Management Plan Template*, 2021), with the addition of the Introduction and the Next Steps sections.

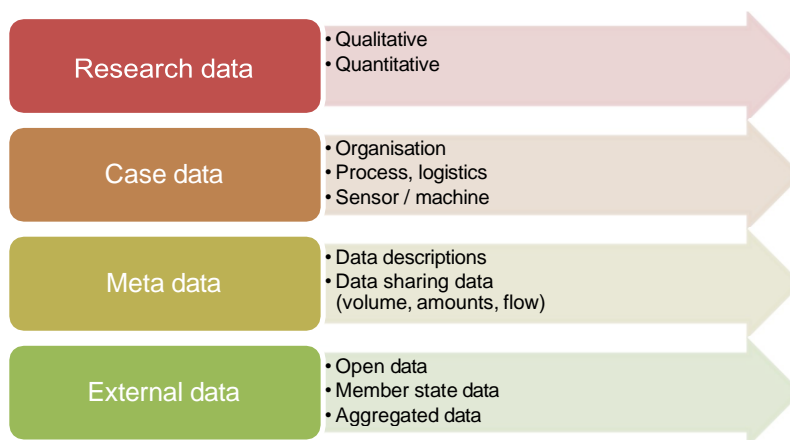
## 2 Data Summary

This section describes four different kinds of data in data categories (Section 2.1) and the way data is re-used in the Data4Food2030 project (Section 2.2). This classification is important because the project does not only deal with research data, but also with data that is collected from the case studies. The data management procedures to be developed during the project are described in Section 2.3.

### 2.1 Data categories

Data4Food2030 will collect data from real-world baseline environments and aggregate it to information & knowledge providing input for the DE4FS as well as other data in several types such as numerical data, texts, images, tables and other formats that are described in Section 3.3.

Except for personal data, several categories of data are used for research and as inputs for the Monitoring System (Figure 1), also referred to as the 'Monitor'. Each data category is described in the following sub sections.



*Figure 1 Input data categories and examples*

For the purpose of the project, for example to create dashboard for the Monitor (<https://monitor.data4food2030.eu/>), it may require combining data sources. This version of the DMP provided guidelines for data management in line with F.A.I.R principles (see Appendix C). The following subsections will describe the four data categories in more detail.

#### 2.1.1 Research data

The research data in this project is quantitative and qualitative data originating from the project activities such as surveys, focus groups, stakeholder meetings and interviews. Management of research data will be done in accordance with the related soft law instruments governing scientific research, such as the European Code of Conduct for Research Integrity, the Guidelines to rules on Open Access to Scientific Publications & Open Access to Research Data in Horizon Europe, and the Guidelines on Data Management in Horizon Europe.

### 2.1.2 Case data

During the project it is estimated that the data will be produced specifically from the case studies. The produced data will be enhanced and combined with the existing data acquired from existing repositories (like the GS1 Global Product Classification), GIS/EO data, unique identifiers (GGN, GLN) for being able to analyse and aggregate the data as applicable. This will require special attention with respect to the anonymisation of data and limitation of public access to competition-critical data. However, for wider publication, data will be abstracted in a way that core messages can be communicated, not revealing trade secrets.

Due to the sensitive data that will be managed during the project, it is scheduled to request a total or partial opting-out of some specific data as they will be incompatible with the need for confidentiality in connection with security issues and with existing rules concerning the protection of personal data (GDPR, NIS Directive) and trade secrets. For this, please refer to the Ethics aspects of the data management in Section 6.

### 2.1.3 Meta data

Meta data are defined as data that describe one or more aspects of the data. This project expects to collect data from the cases and other stakeholders, about the way in which data is shared. For example, the size of data sets or data flows, how many times data is shared, etc.

In addition to that, other data categories, like research data, need to be described with meta data. Per data set, keywords will be provided by the data collector to optimize the possibility for discovery and potential re-use in such a way that it can be harvested and indexed. Section 3.1 describes the usage of for research data, in relation to the open repository in more detail.

Anonymised or pseudonymised data that will be made public, needs to still be compliant with GDPR. It should not be possible to re-identify data to personal data (Gambis et al., 2014). This way to best avoid risk of re-identification of (meta)data will be further investigated during the project. but that depends on how each stakeholder of the project comes to agreement on this.

### 2.1.4 External data

Open data sets that are accessible to the public and can also be used in this project. These can be data about the food system from Member States in an aggregated form or any other open data source that would add value to the Monitor (for example weather data, satellite data, data economy data, food system market data, etc). External, open data does not require data management in the sense of privacy or sensitivity and is therefore only mentioned in this section in the DMP.

### 2.1.5 Personal data

Personal data (name, address, email, telephone, affiliation, gender, etc) are only collected for research purposes, for example to plan interviews and send surveys, not to be part of the Monitor. Please refer to section 6 *Ethics* for more information on how GDPR and Ethics guidelines are applied in the project, as well as an initial Consent Form procedure.

## 2.2 Data re-use

Data will be re-used to gain insights on the usage and exchange of data from one part of the food system to another (frequency, audience, accessibility, value, etc.) to better understand the data economy in the food system. For example, it could be part of the Monitor to understand how many data exchange connections a case study manages (within a value or food loss chain), at the micro level.

Data will also be re-used for analysis of the data at the micro level to aggregate them to a higher, meso or macro level into information.

Data that is personal or has business or competitive value will not be re-used, unless explicitly agreed upon. This could be the case when companies (not part of the consortium), need keep this information private to take advantage of it or keep their position in the food system.

### 2.2.1 Purpose

The data re-used and generated in this project have their purpose in their function as part of the food system. For example, data captured from a farmer's field or at a later point in the food value chain may be transformed into information or knowledge in another part of the food system (e.g., for the consumer). The collection of data for this project has a particular additional objective to capture data to monitor the data economy in the food system, as mentioned before. This project expects to collect data from the cases and other stakeholders, about the way in which data is shared. For example, the volume of data sets or data flows, how many times data is shared, etc. This meta data is anonymised and could be made public, but that depends on of and how each stakeholder of the project comes to agreement on this.

### 2.2.2 Data described

This project will create multiple research data sets as well as re-use data from the stakeholder network. This section describes these data sets in general.

Table 1 captures key elements of the data from each category in section 2.1 to show if it is existing data that is re-used or data that is generated, the origin and utility. The expected size will be estimated in the Work Package and Case Study DMPs. Also, the data utility will be made more specific.

Two more columns have been added to Table 1; "Open / private" and "Level" (micro, meso and macro) which describe if the data in the categories is expected to be openly available or not after its usage. Private data that are not to be shared, may come from organisations in the food system that are related to financial, competitive or trade confidentiality. This project is multi-beneficiary and is it possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement. Whether this is a legal or contractual limitation, should be looked at in more detail per Work Package, Case Study or data set under the supervision of the Work Package lead and when necessary.

The owner of the data will have a veto to ensure that no industrial confidential data will be released to the public. The Data Management officer will suggest methods of anonymizing or pseudonymizing data for each case to ensure that as much data as possible can be published for open access.

*Table 1 Data collection characteristics per category*

Data type	Open / private	Level	Generated / re-used	Origin / provenance	Data utility
Quantitative	Open / private research data	Micro	Generated	Case study organisations and individuals, stakeholders	Describing and monitoring the DE4FS: research in the form of surveys involving stakeholders.
Qualitative	Open / private research data	Micro, Meso	Generated	Case study organisations and individuals, stakeholders	Describing and analysing the DE4FS: research involving focus groups, workshops involving stakeholders.
Case data	Private	Micro	Re-used	Case study systems (ERP, FMS, etc), process, sensors, machines.	Analysing and monitoring the DE4FS, input to Monitoring System.
Meta data	Open	Micro, meso, macro	Re-used	Information about data, data exchange and usage from any source.	Analysing and monitoring the DE4FS, input to Monitoring System.
External data	Private / open	Meso, macro	Re-used	Aggregated data, for example member state or EU level.	Analysing and monitoring the DE4FS, input to Monitoring System.

Any specifics about the data collected will be described per Work Package, Case Study or data set, such as if an embargo is applied to seek protection of intellectual property, the term of storage and any restrictions on usage. If embargo is applied, it will be valid until the patent is applied for / the innovation has been disclosed to the public. If this is not done within a year (public disclosure, patent filing) after discovery, the embargo can be lifted.

At the point of writing this DMP, it is expected that the project will not have to make use of other national, sectorial or departmental procedures for data management. The European Commission, as the funder of the project, offers several guidelines and templates to be used for data management, which are referred to throughout this DMP. The project will also rely on support from the data team at WUR, the project coordinator, as they have experience in for example, the usage of a data repository.

## 2.3 Data management procedures

This project will provide the following data management processes.

1. project data management plan;
2. publishing (open) data;
3. data processing agreements for non-research data exchange and;
4. consent form procedure.

These will be communicated clearly to the entire consortium and Work Package leaders will be required to monitor them in their Work Package.

1. Based on the DMP at project level, each Work Package, case study and/ or if necessary, can per data set. Research data and other data sets that have been uses will be added to the final version of the project DMP or be part of Periodic / Final Project Report. Mainly WP2 that is in charge of the Monitor has to do with the collection of (external) data to be used for the data economy dashboards.

A Data Management Actions form (Appendix E) is provided as a checklist to support structured, GDPR-compliant data management within each Work Package (WP) of the Data4Food2030 project. It ensures that for every research activity involving data collection, such as interviews, surveys, focus groups, or other methods, a record is kept of how data is handled. At the end of the project, the final version of the DMP will provide an overview.

In the final reporting period, each WP lead asked to complete the table for relevant activities, to be included in the periodic reporting, capturing the following key elements:

- The type of data collected (e.g. qualitative, quantitative, metadata, external data)
- Whether the data is open, closed, or private
- Whether data is generated or re-used
- Details on the origin and provenance (e.g. case study participants, systems, sensors, public statistics)
- The source (external party or company) and the intended use of the data

This process improves transparency, traceability, and regulatory alignment across the project, while laying a foundation for responsible data stewardship. The checklist also supports ethical decision-making and open science ambitions by distinguishing between types of data (e.g. scientific publications vs. project deliverables) and ensuring appropriate access and protection levels are in place.

2. To maintain clarity and accessibility, formal project deliverables are published on the designated project website, while scientific publications, datasets, and related outputs are hosted on Zenodo. This distinction supports consistent communication, eases navigation for stakeholders, and strengthens the visibility of validated project milestones. Involving WP6, the procedure for researchers to publish their data with Zenodo follows this workflow (Figure 2). For details, please refer to Section 3.2.

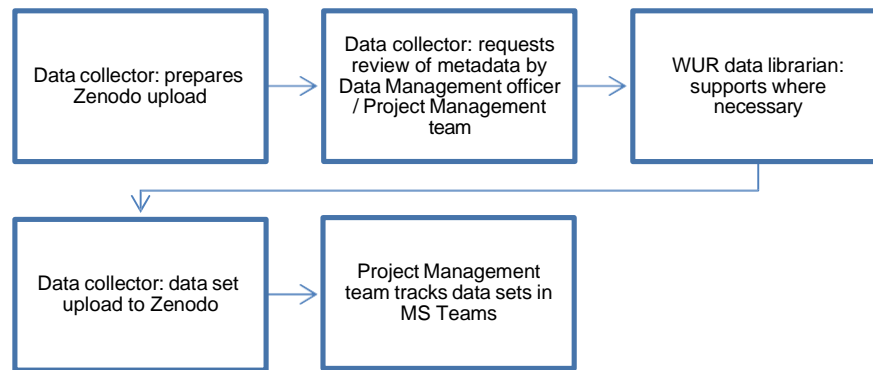


Figure 2 Research data publishing to Zenodo

This procedure can make use of the parts of this DMP about how to create metadata, a README file, anonymisation / pseudonymisation, and which steps to take to upload to Zenodo.

The Data4Food2030 website has been updated with a Zenodo page that holds the first project resources (Figure 3). Forthcoming reports, articles and other types of project resources will be added to this linked Zenodo page (Figure 4) throughout the remainder of the Data4Food2030 project.

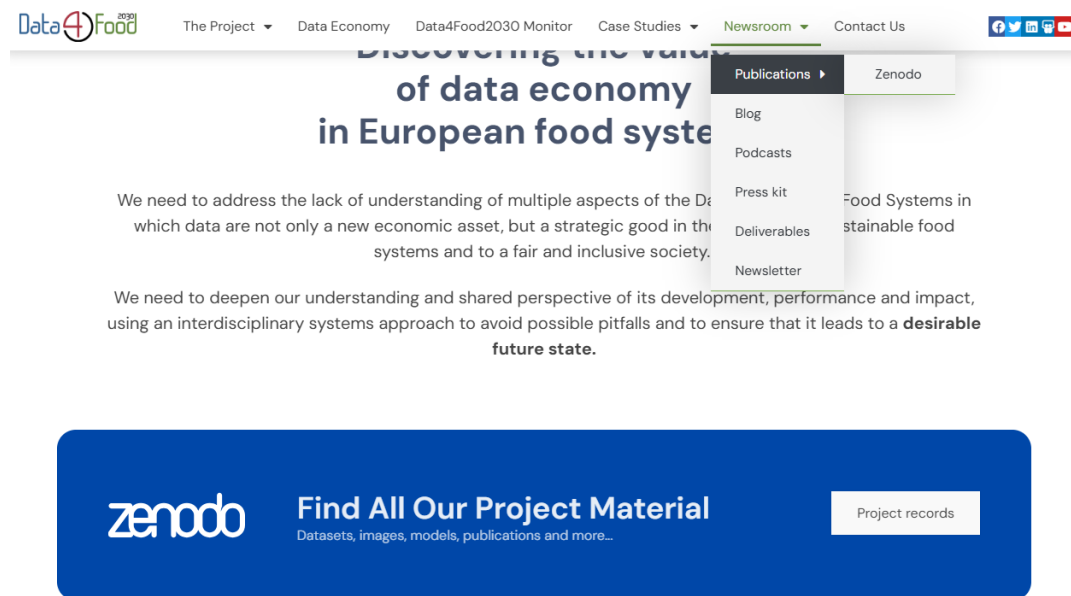
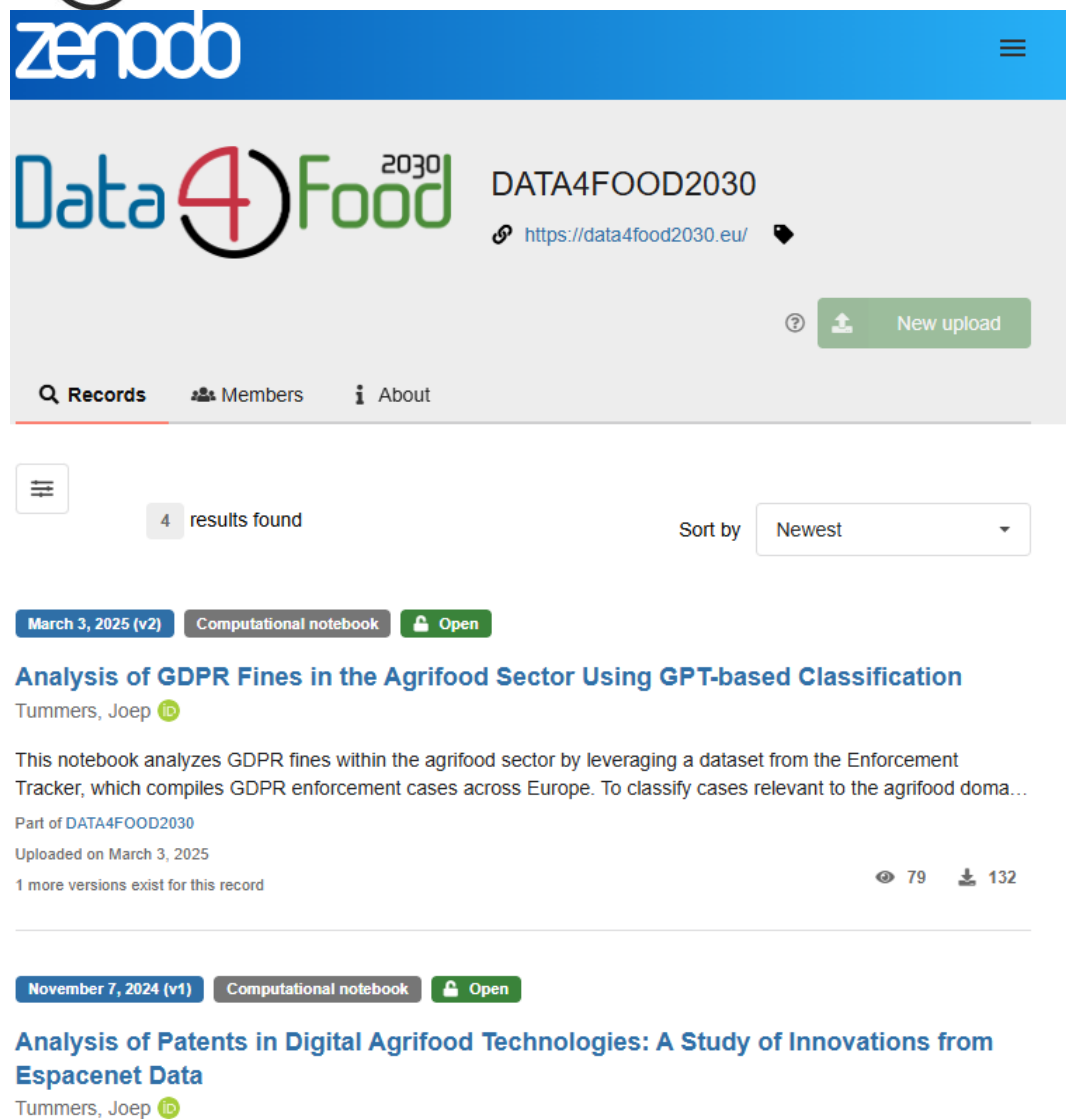


Figure 3 Zenodo access on the Data4food2030 website




zenodo



Data4Food<sup>2030</sup> DATA4FOOD2030  
<https://data4food2030.eu/>

Records Members About

4 results found Sort by Newest

March 3, 2025 (v2) Computational notebook Open

**Analysis of GDPR Fines in the Agrifood Sector Using GPT-based Classification**  
Tummers, Joep 

This notebook analyzes GDPR fines within the agrifood sector by leveraging a dataset from the Enforcement Tracker, which compiles GDPR enforcement cases across Europe. To classify cases relevant to the agrifood doma...  
Part of [DATA4FOOD2030](#)  
Uploaded on March 3, 2025  
1 more versions exist for this record  79  132

November 7, 2024 (v1) Computational notebook Open


**Analysis of Patents in Digital Agrifood Technologies: A Study of Innovations from Espacenet Data**  
Tummers, Joep 

Figure 4 Zenodo page for Data4Food2030 resources

3. Data sharing agreements are drafted and created between the consortium partners at the case level (open/private data). They are quite context specific, because the agreements depend highly on the context, scope and sector of each case. The Common European Agricultural Data Space (CEADS) - <https://ceads.eu> aims to provide model contractual agreements to facilitate data sharing. However, they will be available only after the Data4food2030 project ends. As the cases continue to develop and participate in the DE4FS standard agreements are expected to provide guidance to stakeholders and may be used as a basis for legal data sharing agreements in the cases or for any other Data Sharing Initiatives in the DE4FS for food systems.
4. The consent form guidelines in this DMP are described in a separate easy to follow for researchers and other project team members involved, with WP4 and WP5 (see Appendix B: Informed consent procedure).

### 3 FAIR data

This section describes how the project ensures data is FAIR (Section 3.1), accessible (Section 3.2), interoperable (Section 3.3) and how to increase the re-use of research data (Section 3.4). The way these are organized, depends on the way the data is stored. Please refer to Appendix C for a checklist of FAIR management guidelines for consortium partners.

#### 3.1 Making data findable

Repositories are services that preserve data safely and make them findable. The findability of the data will be managed through the following actions:

1. Data will have Persistent identifiers (PIDs), (Digital Object Identifiers (DOIs)) which are important because they unambiguously identify the data and facilitate data citation and the data will be deposited in trusted repositories such as Zenodo, which assigns DOIs.
2. Data will provide adequate metadata (a description of the key attributes and properties of each dataset) to enable the finding, understanding, citation and reusability of the data. Metadata will provide important context for the interpretation of the data and make it easier for machines to conduct automated analysis. Metadata standard schemes will be followed. Indicatively the following schemes are presented for general data: Dublin Core, CERIF, DDI. For ecology disciplines, EML, ISO 19115 and FGDC-CSDGM (for describing geospatial information), If agricultural information needs to be made findable, the AIMS and AgMES schemes are available. Other types of data (genomics, biological and biomedical data, clinical and non- clinical data) are not expected to be used in the project.

The generated data will be well-documented, and they will have clear license and provenance information. Concerning the documentation, a README file for ensuring that the data can be correctly interpreted and re-analysed by others will be created. This file will include:

- a) a short description of what data it includes along with tables, figures, or sections;
- b) for tabular data: definitions of column headings and row labels, data codes (including missing data) and measurement units;
- c) any data processing steps that may affect interpretation of results;
- d) a description of what associated datasets are stored elsewhere, if applicable and
- e) contact information.

A README file is a document that accompanies the data and provides an explanation of the data, making the data understandable and reusable. Additionally, it contains , i.e., data about the data (see below). A template and an explanation for making a README file is available at WUR (Explanation of the Readme Template for Datasets, n.d.).

When added during deposition of data in a repository, metadata enhances the findability of data. Besides being added to the README file (see above), metadata will be used to describe data in the data repository, e.g., title, authors, keywords, etc. The data repository itself does not check this automatically and this is the responsibility of the data collector. The Data4Food2030 project is supported by a data librarian at WUR to check for necessary and the README file. This support is, however, not limitless and the project team will apply their own data publication process and ask support in the beginning of the project and for specific questions.

## 3.2 Making data accessible

Making use of repositories also enable data accessibility. Data may be archived publicly or with restricted/closed access. Data that can be shared publicly will be made available in the widest form possible and deposited in an Open Data repository. Using the platform of re3data (<https://www.re3data.org/>) several repositories were considered, such as Zenodo, DRYAD, Harvard Dataverse (for general data). Zenodo has been selected to be used during the project for open data publishing, because this platform is also supported by WUR to support with depositing data (Zenodo, DANS-EASY, 4TU.ResearchData) and because Data4Food2030 does not require a domain or topic specific repository (as applicable in for example genealogy or archaeology research).

Zenodo is a repository with a general purpose that enables researchers, scientists, projects and institutions to share, preserve and showcase multidisciplinary research results (data, software and publications). Using Zenodo makes the data accessible, because it is not part of the existing institutional or subject-based repositories of the research communities. It is stored in the CERN data centre in a trustworthy manner and provides access to everyone (open science). It is being used by more than 50.000 researchers and 3000 communities across the world (*Zenodo - EOSC Marketplace*, n.d.).

Zenodo is used at no additional costs unless more than 50Gb, for larger data sets 4TU.ResearchData can hold data sets up to 100Gb for free, if needed. The project management team is exploring appropriate arrangements for Zenodo and will provide a procedure to the project no later than the first dataset upload to Zenodo. It has been verified that Zenodo ensures that the data is assigned an identifier and that it resolves the identifier to a digital object. No documentation or additional software is needed for accessing a Zenodo data set. Additional repositories could be looked at to widen the scope of publishing project results, possibly national repositories per project partner country.

Data will be stored in a separate data usage tracker that allows to document meta data (e.g., the type of data, origin), its storage place and the parties granted to it. This will also specifically include if the data is intended for publication as open access data. The data and will be stored in Zenodo during the project of 4 years and after, for the duration of an additional 10 years.

The general principles for handling Knowledge and IPR within Data4Food2030 have been settled in the Grant Agreement and Consortium Agreement (CA). These principles are in line with Horizon Europe IPR recommendations. Background and foreground (results) will be clearly identified in detail within the consortium agreement and when applicable, granting of access rights will be clearly specified. The IPR strategy and the exploitation management will be handled as well as in T6.3 Integrated toolset for multiplying impact, which includes an IPR plan.

### 3.3 Making data interoperable

Data is interoperable when accessed and processed from multiple sources without losing meaning and integrating that data for mapping, visualization, and other forms of representation and analysis (*Data Interoperability Collaborative*, n.d.). All data files in the project (raw data, processed data, code etc.) used in data collection, processing, and analysis should be made interoperable whenever possible. The guidelines in this chapter can be added to a procedure or guideline for data collectors in the project, in collaboration with WP2, working on the Monitor and WP3 collaborating with the cases.

The produced data will use common formats and standards and community agreed schemas, controlled vocabularies, keywords, thesauri or ontologies where possible in order to be interoperable and be integrated with other data, applications and workflows. For applications and workflows, Open API specification (*FAQ - OpenAPI Initiative*, n.d.) and the ATLAS project (*Interoperability Network - ATLAS*, n.d.) aim to provide standards for the service layer of interoperability.

For the generation, processing and publishing of data files the following formats will be used according to the data types: Web: RDF(S), OWL, SKOS, Containers: TAR, GZIP, ZIP, Databases: XML, CSV, Geospatial: SHP, DBF, GeoTIFF, NetCDF, Moving images: MOV, MPEG, AVI, MXF, Sounds: WAVE, AIFF, MP3, MXF, Statistics: ASCII, DTA, POR, SAS, SAV, Still images: TIFF, JPEG 2000, PDF, PNG, GIF, BMP, Tabular data: CSV, Text: XML, PDF/A, HTML, ASCII, UTF-8, Web archive: WARC.

Agri-food community standards will be considered to make data interoperable, such as GS1 for food product standard references, RmAgro as a reference model for agri-food value chain. There are many initiatives that develop and publish agri-food data standards. This project aims to adopt and build on the most suitable existing initiatives, such as the Minimal Interoperability Mechanisms (*MIMs - Open & Agile Smart Cities*, n.d.), GO FAIR (*GO FAIR Initiative - GO FAIR*, n.d.) and standards applied in agri-food such as from IoF2020 (Breemer & Soum, 2021) and AIM for an ontology standard (*An Update on the Agriculture Information Model - Demeter*, n.d.).

Where applicable, data will include qualified references to other data (from the project, or datasets from previous research).

### 3.4 Increase data re-use

Referring to access licenses, data will have a clear license to govern the terms of its reuse (e.g., Public Domain, Attribution, Non-commercial, No Derivatives, or other). Data4Food2030 will use the Creative Commons licenses and more specifically, the Attribution (CC BY) license and Creative Commons Zero (CC0) which is for dedicating works to the public domain.

A data set that is added to Zenodo will be classified with the proper license. Project members will be instructed publish data sets with the CCBY license as much as possible to be sure to include references to the project publications. It is possible to upload data sets to Zenodo that do not use an open licence in order to publish it with a DOI. Access is then requested (by email) to the data set collector. However, we will not be using this during the project, because sensitive datasets will not require a DOI reference.

The Consortium Agreement (CA) of the project states that the consortium partners shall, where necessary, conclude a separate data processing, data sharing and/or joint controller agreement before any data processing or data sharing takes place. In this agreement, it needs to be stated if and which data in the project will not be used by Third Parties, even after the project.

Data quality is ensured by verifying the upload to Zenodo (readme file and format), supporting project members in the usage of Zenodo and enforcing the creation of re-usable data sets.

### **3.5 Other research outputs**

Other outputs of the project, such as digital outputs (e.g., software, workflows, protocols, models, etc.) or physical outputs (e.g., new materials, antibodies, reagents, samples, etc.) will be monitored by the project team, by asking for specific other formats during periodic reporting. It is not expected that software will be uploaded as open data, but rather data standards, reference architectures and Minimum Interoperability Mechanisms (MIMs). Also, methodologies for stakeholder engagement, if openly accessible, can be shared through Zenodo.

## 4 Allocation of resources

In general, each Work Package leader is responsible for the data sets that are used for the Monitor and / or for research of the DE4FS during the project. This responsibility can be delegated to the Task leader or researcher responsible for data collection; the data collector. The project management team is available for support. Any changes in this responsibility can be listed in the DMP per research activity.

Table 2 lists the roles responsible for the data management tasks within the project.

Due to centralized support processes within WUR, it is expected that data management will not lead to extra cost for the project. If any costs for data curation and storage/preservation will be incurred, the DMP will be updated, and costs will be processed as material cost as part of Work Package 7 Project Coordination.

*Table 2 Resource allocation for data management*

Data collection	Work Package leader, Case Study leader or individual researcher (data collector)
Data quality	Data collector
Storage and backup	Zenodo and data collector
Data archiving / publishing	Zenodo
Data stewardship / support	WR data team, EU, CESSDA
Data management support and liaison with legal experts	Data Management officer

At this point in the project in month 4, the installation of a data access committee does not seem necessary because responsibilities have been assigned to the roles above.

## 5 Data security

WUR provides an overview of permitted storage options for research and non-research data storage. The most suitable storage location depends on factors such as data type, size, retention, access rights, etc. Using a new Data Storage Finder, researchers can find the most adequate storage location for their data (*Find the Best Data Storage Location* - WUR, n.d.).

For the project, data security is in place using the project collaboration platform MS Teams which is hosted on the WUR network. Using Google drive, for example, is not allowed, due to the risk of security holes and loss of control over the data.

Data storage for non-research data will be performed in a secured form (e.g., data encrypted with a strong cryptographic protocol) in servers indicated by the case studies or the technology providers, and agreed within the consortium, further detailed as part of T3.2 activities.

Research data stored in Zenodo is securely stored on CERN premises, is being backed up and checked for data compatibility.

## 6 Ethics

Data4Food2030 is fully compliant with the Do No Significant Harm (DNSH) Principle since its activities (both those carried out during the project as well as the expected life cycle impact of the innovations at a commercialization stage, including both products and services), substantially contribute in a positive way and/or do not significantly harm any of the six environmental objectives set out in Articles 9 and 17 of the EU Sustainable Finance Taxonomy Regulation (EP, 2020).

The EU code of conduct on agricultural data-sharing constitutes a document that looks at general principles for sharing agricultural data within the agri-food chain. The consortium will follow this non-binding code that sheds greater light on contractual relations and provide guidance on the use of agricultural data (COPA & COGECA, 2020).

When processing personal data, the consortium will comply with the GDPR principles. This means that at the outset, when the data cannot be anonymized fully, in the development of the project, Article 5(1) of the GDPR, enshrining the principles of personal data protection, will be considered.

Involvement of stakeholders, end users and the linked feedback and participatory processes are very important within this project. This includes aspects where ethical and gender issues must be checked (e.g., the selection of stakeholders, their consent for participation in the experiments, data privacy and protection). In line with the ethics self-assessment, actions will be executed for involvement of humans and for the protection of personal data.

The consortium will comply with the General Data Protection Regulation (GDPR) (UE) 2016/67. The GDPR has been analysed and reviewed in order to align the project objectives accordingly. All partners should have their internal procedures updated and are compliant with current legislation. To be sure that partners comply, a compliance check may be organised asking the partners to confirm compliance.

### 6.1 Informed Consent Procedure

The processing of personal data is generally prohibited or restricted unless explicitly permitted by the law. The legal bases for processing personal data are described in the GDPR article 6. One of these legal bases is the use of freely given and informed consent. Due to the nature of Horizon Europe ethical guidelines, it is generally recommended as a necessity that participants from whom data is collected, informed consent procedures are used.

Informed consent is described in Article 7 of GDPR and specified in recital 32 of the GDPR. Consent must be freely given, specific, informed and unambiguous. By freely given, this means consent can also be freely rescinded.

Stakeholders who are directly involved in the project will provide research data with the project consortium under the necessity to perform project activities as a contractual obligation. This is since consent from these stakeholders, particularly the case studies cannot freely give consent as they already have a legal basis for engaging in the project. Nevertheless, the data of all stakeholders shall be handled with confidentiality and care by all partners in the consortium and only for the purposes of necessity. Consent forms may be used where necessary and the consent form procedure that will be drafted will indicate this more clearly.

To obtain informed and freely given consent, the data subjects sharing their data must be informed of the identity of the person collecting the data. The data subject shall also be informed how the data is processed as part of the consent form. It is necessary to inform the user which categories of data can be collected, to which purposes this data is collected, how long it will be retained for,

and where the data will be used. If necessary, a separate consent form procedure will be created for images that are taken as well as the usage of these images (personal data identifying the users) to disseminate the project.

A template for an informed consent form is available for all project partners (see Information Letter in Appendix A: Data4Food2030 Information letter V1.1). When conducting an interview, survey, focus group, workshop, or another activity where informed consent is needed, the template is used and adjusted according to the needs of the activity. The informed consent form can either be presented in an online format via MS forms or printed and signed by the person involved in the research activity. An instruction on how to use the online format via MS forms is available in

Appendix B: Informed consent procedure.

The data retention period of data that is collected through these activities is the duration of the project (4 years), plus 10 years.

## 6.2 Stakeholder groups

Once different types of stakeholders have been identified (Task 4.1), the data collection can be described per stakeholder group. For data of vulnerable groups, consent is needed to process their data. If consent is the legal basis on which we will acquire data, this condition will be met.

## 6.3 Gender-inclusive approach and integration

Data4Food2030 acknowledges the inherent gender disparities in economies and is committed to integrating the gender dimension into our data management to foster inclusivity and prevent gender-based inequalities. When necessary for the purposes of a research activity, we will collect data across gender, following SAGER publication guidelines. Moreover, we will promote gender balance in research, decision-making, and advisory boards, aligning with "Horizon Europe" requirements to empower women in research and innovation. Furthermore, we will ensure equal data access for woman, men, and gender-diverse individuals. Through these efforts, Data4Food2030 contributes to a fair and inclusive data economy for food systems.

## 6.4 Data consent retraction procedure

The consent forms shall include information on which e-mail to contact to access, rectify, delete and/or retract their data. In the case of consent being rescinded the processing of the user's data is to stop immediately. The retraction of consent should be completed as easy as consent was acquired. The WP lead in charge of the data collection or the WP7 team should confirm that data processing has stopped. If a box is ticked to obtain consent, then a box can be unticked for processing to stop.

In the case of data being deleted, the WP team should write to the user and ask for confirmation of how their data was originally acquired. If during this project it is likely that multiple partners may have access to the data of the user, the controller may wish to check with other partners to see if they hold data on the user wishing to have their data deleted (or verified or rectified). In the case of data deletion, the data of the user must be deleted within 30 days.

## 7 Next steps

The next and final version of the Data Management Plan (DMP) will be developed according to the deliverable deadlines. An overview of expected changes is displayed below in Figure 3.

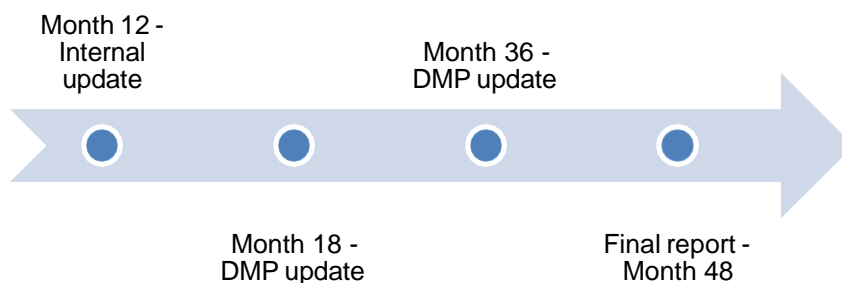


Figure 5 Data Management Plan development timeline

In more detail, the data management activities throughout the project are expected to be necessary in the time frames as listed in Table 3. The final version of the DMP will be ready by month 48.

Table 3 Data management activities during project

Time frame	DMP activity
Month 12	Data Management Plan Process description Data publishing procedure (Zenodo)
Month 18	Informed Consent template and procedure Check for other research outputs (no other usable research outputs so far)
Month 36	Re-check for other research outputs GDPR compliance and of re-identification risk checklist FAIR data management and checklist
Month 48	Overview of data collections per WP/partner (instead of DMPs per Work Collection of GDPR compliance and of re-identification risk checklists Package / Case Study (if necessary) – these are not included) Recommendations and lessons learned for future projects

## 8 References

- An update on the Agriculture Information Model - Demeter.* (n.d.). Retrieved December 23, 2022, from <https://h2020-demeter.eu/an-update-on-the-agriculture-information-model/>
- Breemer, A., & Soum, C. (2021). *D3.5 Guidelines for the use of IoT related standards in smart farming and food security.* [www.iof2020.eu](http://www.iof2020.eu)
- CESSDA Training Team. (2022). *CESSDA Data Management Expert Guide.* <https://dmeg.cessda.eu/>
- COPA & COGECA. (2020). *EU Code of conduct on agricultural data sharing by contractual agreement.*
- Data Interoperability Collaborative.* (n.d.). Retrieved December 23, 2022, from <https://www.data4sdgs.org/initiatives/data-interoperability-collaborative#:~:text=Interoperability%20is%20the%20ability%20to,structure%20and%20content%20of%20datasets.>
- Explanation of the readme template for datasets.* (n.d.). Retrieved December 19, 2022, from <https://library.wur.nl/WebQuery/edepot/534144>
- FAQ - OpenAPI Initiative.* (n.d.). Retrieved December 23, 2022, from <https://www.openapis.org/faq>
- Find the best data storage location - WUR.* (n.d.). Retrieved December 19, 2022, from <https://www.wur.nl/en/show/Find-the-best-data-storage-location.htm>
- GAIA-X - Home.* (n.d.). Retrieved December 21, 2022, from <https://www.data-infrastructure.eu/GAIA-X/Navigation/EN/Home/home.html>
- Gambis, S., Killijian, M. O., & Núñez Del Prado Cortez, M. (2014). De-anonymization attack on geolocated data. *Journal of Computer and System Sciences*, 80(8), 1597–1614. <https://doi.org/10.1016/j.jcss.2014.04.024>
- GO FAIR Initiative - GO FAIR.* (n.d.). Retrieved December 23, 2022, from <https://www.go-fair.org/go-fair-initiative/>
- Horizon Europe Data Management Plan Template.* (2021).
- Interoperability Network - ATLAS.* (n.d.). Retrieved December 23, 2022, from <https://www.atlas-h2020.eu/interoperability-network/>
- MIMs - Open & Agile Smart Cities.* (n.d.). Retrieved December 23, 2022, from <https://oascities.org/minimal-interoperability-mechanisms/>
- Zenodo - EOSC Marketplace.* (n.d.). Retrieved December 21, 2022, from <https://marketplace.eosc-portal.eu/datasources/eosc.openaire.0a02f13310296033694acead588a773b>

## Appendix A: Data4Food2030 Information letter V1.1

Please take some time to read this information and ask questions if anything is unclear. Contact details can be found at the end of this document.

### Project background and purpose

The aim of the [Data4Food2030](#) project is to define and map the development, performance and impact of the data economy for food systems (DE4FS) in order to create new insights and opportunities related to the digital transformation of food systems. Data and data-driven innovations are already shaping food systems and the ways we produce, consume and govern food, and the project aims also at reducing the existing knowledge gap about these data processes and their impacts, including on different food system elements and stakeholders.

Together with (the input from) various stakeholders the researchers want to identify drivers and barriers related to the development of a data economy for food systems, and turn these into opportunities, recommendations and solutions for the development of a data economy.

### Organisation

Wageningen Economic Research is coordinating the Data4Food2030 project and the research for this study is being undertaken by the [consortium partners](#). This *<interview | survey | focus group | workshop | other >* is part of the activity *< activity name >*, undertaken by *< activity partners >*.

This project is funded through the EU commission through the Horizon Europe funding program (grant agreement No.101059473). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

### Participation in the study

Participation in this research is entirely voluntary, and you are encouraged to seek clarification or ask any questions before deciding to take part. If you agree to participate, you will be asked to sign this consent form. This research is dedicated to include all genders and to ensure a diverse perspective through gender balance.

If you agree to take part in *<interview | survey | focus group | workshop | other >*, *<Short sentence about contents of the study>*.

*< Two sentences about why participation in this study is important >*

*<If applicable, mention the compensation that the participant will obtain for taking part>*

### How will my data be used?

The information from the *<interview | survey | focus group | workshop | other >* will be used by the researchers in the Data4Food2030 project for a report to the European Commission, to develop infographics that will be publicly available through the website of the Data4Food2030 project (<https://data4food2030.eu>) and for open access scientific article(s).

*<IF personal data is stored: Personal data (name, organization, e-mail, phone number) is collected for the purpose of contact, follow-up and invitation to this study only and will be anonymised in any reports,>*

infographics or publications. >

<IF special categories of personal data are stored: (gender) is collected for the purpose of being able to prove that Data4Food2030 is committed to integrating the gender dimension into our data management to foster inclusivity and prevent gender-based inequalities. Moreover, we will promote gender balance in research, decision-making, and advisory boards, aligning with "Horizon Europe" requirements to empower women in research and innovation. Furthermore, we will ensure equal data access for woman, men and gender-diverse individuals. In this case additional security measures will take place>

<IF personal is not stored: No personal data (name, organization, e-mail, phone number) is collected in this study>

The information will be processed in accordance with the provisions of the GDPR and will be deleted no later than 10 years after the project ends. Your personal data will not be shared with other parties other than the consortium partners involved in this project.

The information will be retained by <Organization> and will only be used for the purpose of research. The security of personal data and by the processors takes place on the basis of generally accepted standards and best practices. Please refer to *the* <Link to organization information security policy: e.g.: [WUR information security policy](#)> and the <Link to organization data policy: e.g.: [WUR data policy](#)> for more details.

By supplying this information, you give consent to storing your information for the purposes and uses outlined above. Any use of the information beyond the scope or duration of this project will require the researchers to contact you for (re-newed) consent.

There are no known risks in taking part in this study.

### Research results

You may request a summary of the research findings by contacting the task leader, < Name of task leader> of the <Organization>: <Insert email of task leader>. At any time, you are free to withdraw consent, have personal data removed or adjusted and to access to the personal data have been processed by contacting the task leader.

### Concerns or complaints

If you have any concerns about the project, your involvement in it or this consent form, please discuss this with the researcher undertaking the <interview | survey | focus group | workshop | other > in order to find out how your concern will be addressed.

<!Delete this: Note to researcher: Ensure that there is gender balance among the participants in your activity. !>

---

### Consent:

- I have read the information presented in this information letter. I have had the opportunity to ask any questions related to this research and received satisfactory answers to my questions.
- I understand that relevant sections of the data collected during the research may be accessed by individuals from the Data4Food2030 project.

- With full knowledge of all the foregoing, I agree that my answers will be processed as part of the Data4Food2030 project.
- *<If applicable: I am aware that excerpts from the discussion may be included in publications to come from this research. Quotations will be kept anonymous. >*

**Yes**

**No**

**Name:**

**Date:**

**Signature:**

*<If applicable: I agree to be contacted again by the researchers for clarification or elaboration on my input in the discussion. >*

## Appendix B: Informed consent procedure

Using the MS Forms template:

1. Open this link: [Link to template](#)
2. Press the button next to: "Duplicate this form to use as your own"



Duplicate this form to use as your own.

Duplicate it

3. Log in with your company Microsoft Account.
4. Edit the information letter according to your study needs. The fields that need to be changed are indicated in different <colors>. Indicate what type of study you are conducting (interview, survey, focus group, workshop, other) and fill in the required information.
5. You can now start making your form by making a new section.
6. After the study: Export the study results and store them in a safe location, according to the data management plan.
7. Good luck with your form! Don't forget to take into account the gender balance among the study participants.

Using a printed version:

1. Edit the information letter according to your study needs. The fields that need to be changed are indicated with grey <highlights >. Indicate what type of study you are conducting (interview, survey, focus group, workshop, other) and fill in the required information.
2. Print the information letter and make sure the study participants read and sign the form before the study.
3. After the study: scan the signed forms and make sure to store them as a .pdf in a safe location according to the data management plan.
4. Destroy the paper copy after storing the scanned file in a safe location according to the data management plan.
5. Good luck with your form! Don't forget to take into account the gender balance among the study participants.

## Appendix C: F.A.I.R. Data management guidelines

FAIR Principle	Checklist Item (Agri-Food Focus)	Check (✓ / ✗ / Notes)
<b>Findable</b>	Persistent identifiers (e.g. DOIs) are assigned to datasets related to farming, food systems, or trials	
	Metadata includes detailed descriptors such as crop type, location, input methods, and season	
	Metadata explicitly includes the dataset's unique identifier	
	(Meta)data is registered in agri-food repositories (e.g. GODAN, AgInfra, FAO databases) for discoverability	
<b>Accessible</b>	Open, standardized protocols (e.g. HTTP, FTP) are used to enable dataset retrieval	
	Protocols used are open, free, and universally implementable	
	Authentication and authorization procedures are in place where access needs restriction	
	Metadata remains accessible even when underlying data is no longer available	
<b>Interoperable</b>	Machine-readable formats (e.g. CSV, JSON-LD, SDMX-Agri) compatible with agri-food platforms are used	
	Controlled vocabularies or ontologies (e.g. Agrovoc, FoodEx2) are applied	
	Datasets include qualified references to related data sources (e.g. soil, weather, supply chain)	
<b>Reusable</b>	Datasets are described with domain-relevant attributes (e.g. fertiliser use, water inputs, feed composition)	
	Data usage is governed by clear, accessible licenses (e.g. Creative Commons) reflecting reuse conditions	
	Documentation includes data provenance: who collected the data, when, where, and how	
	Datasets comply with relevant standards (e.g. INSPIRE for geodata, ISO 22005 for traceability)	

## Appendix D: GDPR Compliance & Re-identification Risk Checklist

Extended checklist that integrates GDPR-focused risk mitigation with Re-identification Risk. All Data4food2030 consortium partners can self-assess these and mark or add notes. This checklist is provided as a service to the consortium, each partner is responsible for their own privacy policy enforcement.

Category	Checkpoint	Check (✓ / ✗ / Notes)
<b>Legal &amp; Organisational</b>	Up-to-date privacy policy compliant with GDPR	
	Appointed a Data Protection Officer (DPO), if required	
	Internal data protection and security policies in place	
	Data Processing Agreements (DPAs) signed where relevant	
	Documented anonymisation procedures and responsibilities	
<b>Data Processing &amp; Minimisation</b>	Overview of what (personal or sensitive) data is processed	
	Legal basis for processing defined (e.g. consent, legitimate interest)	
	Procedures to uphold data subject rights (access, erasure, objection, etc.)	
	Applies data minimisation and purpose limitation	
	Includes only strictly necessary data for the intended purpose	
	Applies proven anonymisation/pseudonymisation techniques	
	Removes or aggregates high-risk identifiers (e.g. postal codes, timestamps, rare traits)	
	Assesses re-identification risk using recognised models (e.g. k-anonymity, l-diversity)	
	Aligns risk thresholds and responsibilities across all stakeholders	
<b>Metadata Review</b>	Checks metadata for indirect identifiers (e.g. contributor IDs, geolocation, timestamps)	
	Anonymises file properties and naming conventions that may contain identifiers	

<b>Stakeholder Consultation</b>	Project stakeholders agree on acceptable levels of residual risk	
	Governance roles clearly defined for approving data publication	
	Procedures in place for dataset updates or withdrawal if risks change	
<b>Technical &amp; Organisational Measures (TOMs)</b>	Appropriate security measures in place (e.g. encryption, access control)	
	Risk assessments or DPIAs conducted where applicable	
	Data breach response plan available	
	Physical access controls in place	
	Access/licensing terms apply when full anonymisation isn't possible	
	Monitoring for re-identification or misuse post-publication	
	Contact point established for data subjects or third-party concerns	
<b>Collaboration &amp; Transparency</b>	Roles clarified (controller, processor, joint controller)	
	International transfers avoided unless protected by valid safeguards (e.g. SCCs)	
	(Sub-)processors identified and disclosed	
	Joint controllership agreements documented where relevant	
	All stakeholders involved in reviewing and approving data publication	
	Public datasets include GDPR-compliant data sharing statements	

## Appendix E: Data management Actions per Data4Food2030 Work Package

WP: <number>

WP lead: <name>

Please use the table below for each research activity that required data collection.

<b>Research activity</b>	<b>Interview, survey, focus group, other (please add)</b>
<b>Data type (see Table 1 below)</b>	<input type="checkbox"/> Qualitative <input type="checkbox"/> Quantitative <input type="checkbox"/> Case data <input type="checkbox"/> Meta data <input type="checkbox"/> External data
<b>Open / private</b>	<input type="checkbox"/> Open research data <input type="checkbox"/> Closed research data <input type="checkbox"/> Open and closed research data <input type="checkbox"/> Private data <input type="checkbox"/> Open data
<b>Generated / re-used</b>	Generated / re-used
<b>Data origin / provenance</b>	<input type="checkbox"/> Case study organisations and individuals <input type="checkbox"/> Stakeholders involved in the case study <input type="checkbox"/> Systems (e.g. ERP, FMS) used in the case study <input type="checkbox"/> Processes, sensors, and machines involved <input type="checkbox"/> Information about data sources, data exchange, and usage <input type="checkbox"/> Aggregated data (e.g. national or EU-level statistics) <input type="checkbox"/> Other (please add below): - ... - ...
<b>Data source</b>	External / Company
<b>Data utility</b>	<describe what the data is used for>