

Data Management Plan Version 1



**Eco
Ready**

**Ioannis Manikas
Czech University of Life
Sciences - CZU Prague**



Technical References

Project Acronym	ECO-READY
Project Title	Achieving Ecological Resilient Dynamism for the European food system through consumer - driven policies.
Project Coordinator	Czech University of Life Sciences (CZU)
Project Duration	48 months

Deliverable No.	7.2
Dissemination level ¹	Confidential
Work Package	7 Project Management
Task	7.4
Lead beneficiary	CZU
Contributing beneficiary(ies)	CZU
Due date of deliverable	M6
Actual submission date	M6

¹ PU = Public
 PP = Restricted to other programme participants (including the Commission Services)
 RE = Restricted to a group specified by the consortium (including the Commission Services)
 CO = Confidential, only for members of the consortium (including the Commission Services)

Document history

V	Date	Beneficiary	Author
1	16/05/2023	CZU	I.Manikas
2			



Summary

This document contains the first version of the Data Management Plan (DMP) for the ECO-READY RIA HE project, addressing what and how data and metadata will be generated, collected, handled and preserved, in repositories following the FAIR principle (findable, accessible, interoperable and reusable) as well as possible. According to the European Commission, a DMP is a key element of good data management, describing the data management life cycle for the data to be collected, processed, and/or generated by a Horizon Europe project.

As part of making research data findable, accessible, interoperable, and reusable (FAIR), a DMP should include several pieces of information. First is the handling of research data during and after the end of the project, and this is particularly relevant for ECO-READY's Knowledge Infrastructure, including the Observatory and its toolset, and the network of Living Labs. Second is what data will be collected, processed, and/or generated, third which methodology & standards will be applied, and fourth whether data will be shared/made open access. Finally, it should contain information about how data will be curated and preserved during and after the project.

The DMP of the ECO-READY project is developed based on the Horizon FAIR Data Management Plan (DMP) template provided by the European Commission and recommended to be used on voluntary basis.

Disclaimer

This publication has been produced within the ECO-READY project which has been funded by the European Union Horizon Europe Research and Innovation Programme under grant agreement n° 101084201. The views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Research Executive Agency. Neither the European Union nor the Research Executive Agency can be held responsible for them.



Table of contents

Technical References	2
Document history	2
Summary	3
Disclaimer	3
Table of content	4
1 Data Summary.....	5
2 FAIR data.....	6
2.1 Making data findable, including provisions for metadata.....	6
2.2. Making data openly accessible.....	6
2.3. Making data interoperable	7
2.4. Increase data re-use (through clarifying licenses).....	7
3. Allocation of resources	7
4. Data security	7
5. Ethical aspects	7



1 Data Summary

The project involves data collection, including personal data and metadata in the context of the stakeholders engagement tasks (WP2), the implementation and monitoring phase of the Living Labs (WP3), as well as metadata from accessing the Observatory database and utilizing the toolset (WP4). All processing of personal data will be conducted in accordance with the provisions of: a) the GDPR (Regulation (EU) 2016/679)¹, b) the Universal Declaration of Human Rights and the Convention 108 for the Protection of Individuals with Regard to Automatic Processing of Personal Data, and c) the national laws applying its provisions, including those governing the acquisition of valid consent. Data managed during the project will be processed only under the following preconditions which need to be met: (i) When the data subject has given her/his consent; (ii) When the processing is necessary for the performance of or the entering into a contract; (iii) When processing is necessary for compliance with a legal obligation; (iv) When processing is necessary in order to protect the vital interests of the data subject. Personal data managed within ECO-READY will be anonymized and stored in a form which does not permit identification of users. ECO-READY will establish a data management framework that guarantees security of collected data from potential abuse, theft, or loss. Research data that will be collected and/or generated by the project will be -in principle- made open and offered to the Open Research Data Pilot, in which ECO-READY declares its intention to participate. This Data Management Plan DMQP (D7.2) details what data the project will generate (i.e. content, type, format, volume), which standards and methodologies will be used for data collection and management, whether and how they will be exploited and how they will be made findable, accessible, interoperable and reusable (FAIR). The first version of the Data Management Plan is submitted by the project leader (CZU), and will be continuously updated until the end of the project. 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. The ECO-READY project partners must process personal data under the Agreement in compliance with the applicable EU, international and national law on data protection (in particular, Regulation 2016/67916). They must ensure that personal data is:

- processed lawfully, fairly and in a transparent manner in relation to the data subjects
- collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes
- adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed
- accurate and, where necessary, kept up to date
- kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data is processed, and
- processed in a manner that ensures appropriate security of the data.

The ECO-READY project partners may grant their personnel access to personal data only if

¹ https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG



it is strictly necessary for implementing, managing and monitoring the Agreement.

ECO-READY will collect and use different types of data. Primary data will include qualitative data derived from interviews with experts, Delphi surveys, consumer surveys, and inputs collected during specific events (stakeholder working group meetings, workshops and capacity building events). Quantitative data will include satellite data and geoinformation on land use and land use change, as well as environmental DNA (eDNA) metabarcoding.

The Observatory database will combine field real-time data deriving from the living labs as well data from high-resolution images (when necessary), UAV orthophotos (hyperspectral or multispectral), and data from third-party sources such as Copernicus, EU Sentinel missions, USA Landsat missions, etc.

Molecular approaches based on environmental DNA (eDNA) metabarcoding have been chosen because they allow to generate large amounts of data, improving current biodiversity assessments and are able to simultaneously determine species distributions, ecologically important patterns of species diversity and detect novel species. We are planning to conduct a meta-analysis based on the Global Biodiversity Information Facility (GBIF) records and checking regional geographical fit with species lists in available literature. Hence, we could compare the Next-Generation Sequencing (NGS) identified species and test their geographical fit by registered coordinates².

The secondary data will be collected as outcome of the Structured Data Review, in order to outline the food social-ecological system at a European level; data will be searched in the scientific literature, in the EU project reports, but also in National projects, in EUROSTAT database and in other sources.

Taking into account the importance of data management, the first version of the data management plan is presented in this document. This document will be updated in month 24 (D7.4) and month 48 (D7.5) as part of Task 7.4. This task will ensure the proper handling of data that is collected within the project. 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 owner of the data, i.e. generally the source of data will have a veto to ensure that no industrial confidential data will be released to the public.

2 FAIR data

2.1 Making data findable, including provisions for metadata

ECO-READY will produce multispectral data in several types as numerical data, texts, images, tables and other formats that are described below from the 10 Living Labs. When relevant, the produced data will be enhanced and combined with established databases (e.g. MAGNET database) models, possible model application sequences in dealing with problems, data requirements, literature reviews and summarized in an open database, managed by the observatory and accessible through an IT-based selection and advisory tool. The findability of the data will be succeeded as:

- The data will have Persistent identifiers (PIDs) (Digital Object Identifiers (DOIs)) which are important because they unambiguously identify the data and facilitate

² https://www.gbif.org/occurrence/map?has_coordinate=true&has_geospatial_issue=false



data citation. Furthermore, when relevant, the data will be deposited in trusted repositories as Zenodo, which assigns DOIs.

- The data will have rich metadata which will support findability, citation and reuse. Rich metadata will provide important context for the interpretation of the data and make it easier for machines to conduct automated analysis. Standard metadata schemes will be followed. Indicatively the following schemes are presented: Dublin Core, CERIF, DDI (general data) and EML (specific for ecology disciplines), ISO 19115 and FGDC-CSDGM (for describing geospatial information), AgMES (Agricultural information).

2.2. Making data openly accessible

The shared data will be deposited in an Open Data repository which will be identified through the platform of re3data (<https://www.re3data.org/>). Indicatively some repositories are mentioned herewith: Zenodo, DRYAD, Harvard Dataverse (for general data) etc. However, 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). Only anonymised and aggregated data will be made open to ensure that data subjects cannot be identified in any reports, publications and/or datasets resulting from the project. The general principles for handling Knowledge and IPR within ECO-READY will be settled in the GA and 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 the WP6.

2.3. Making data interoperable

Data interoperability refers to the ability of systems and services that create, exchange, and use data to have clear, shared expectations for the contents, context and meaning of that data. As mentioned in Section 2.1, ECOREADY will adopt in its data management methodology the use of metadata vocabularies, standards (e.g. Dublin Core Metadata standard) and methods that will increase the interoperability of the data collected/generated through its activities. 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. Indicatively, the data files generated will be processed and submitted to the repositories mentioned above with the following formats according the data type: 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.



2.4. Increase data re-use (through clarifying licenses)

The generated data will have clear license and provenance information. Concerning the documentation, a README file for ensuring that the data can be correctly interpreted and reanalyzed 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; e) contact information. Referring to the license issues, data will have a clear license to govern the terms of its reuse (e.g. Public Domain, Attribution, Non-commercial, No Derivatives, or other). Indicatively, ECO-READY will use the Creative Commons licenses: the Attribution (CC BY) license and Creative Commons Zero (CC0) which is for dedicating works to the public domain. Archiving and storage/preservation: Data storage will be performed in a secured form (e.g. data encrypted with a strong cryptographic protocol) in servers agreed within the consortium.

3. Allocation of resources

The costs to make the ECO-READY datasets FAIR can be evaluated as a part of the costs allocated for the development and implementation of the Observatory toolset. No additional costs will be required for that purpose. The grant funding covers these costs. The tools that are planned in this project, including the mobile app, are designed to make data publicly available and therefore the costs to make data FAIR were included during the proposal. The website of the ECO-READY project will remain active and accessible for stakeholders for at least three years after the end of the project. The potential value in this is that all results are available for download beyond the duration of the project.

4. Data security

All project partners will store the data in a cloud system that guarantees replication between regions and access patterns to ensure high security. CZU as the ECO-READY project Coordinator will store personal data of stakeholders in an internal access management system with next generation firewalls, intrusion protection, etc. for maximum security. Regarding the project website, daily backups will run to ensure that data recovery is possible.

The protection of personal data in the ECO-READY project is fully in line with the EU General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679). Each partner



conducting an activity that requires data collection (i.e. Data Partner³) is responsible for the data handling and data security for the respective dataset. All project partners are responsible for processing data using appropriate means, such as private servers or cloud service providers that adhere to the relevant legal data protection requirements (e.g., GDPR) and will ensure that this data is protected, and any necessary data security controls have been implemented, to minimize the risk of information leak and destruction.

Partners shall not be liable in any case for any noncompliance with or breach of applicable data protection regulation(s) (such as the GDPR), at European Union and national level, incurred by another Partner. The noncompliant Partner shall be held solely and exclusively responsible for the compensation of any damages, including any administrative fines imposed by the competent Supervisory Authority and/or any other potential penalties in accordance with relevant legal provisions (such as Art. 83 GDPR) and/or any compensation claims by data subjects or any other third party.

5. Ethical aspects

All collection and processing of data will happen in strict adherence to all ethical and legal requirements, as described in D7.6 General Ethical Guidelines. This means that informed consent will always be obtained from the participants, a process during which they will be informed of the aim of the study, the use of their data, and their rights, including their right to withdraw from the study. Project specific data protection policy and the contact details of the data protection officer will be provided by appropriate partners to all participants via written information sheets, which will be written in a language understandable to the research subjects.

It may be possible to anonymise some data, but where this is not feasible, data will be pseudonymised to ensure a high level of security when it is stored and processed. To respect the data minimisation principle, only those data categories that will be necessary in light of the objectives of the project will be processed.

Ethical approval will be obtained for all research activities involving human participants, from the appropriate institutional ethics committees. All data collected from human participants will be stored separately from the project data with the only link being through a code which will be accessible to only the lead researcher in each country where that data has been collected. Therefore, the project data will not be fully anonymised until the link is destroyed. All participants will be made aware of this as part of the informed consent process. Participants will also be told for how long their personal details will be retained (10 years), and when and if the code/link will be destroyed. All data will be collected on laptops or handheld devices and then uploaded to a central storage and removed from the portable devices. For activities where audio or video recordings are taken, the data will be

³ Data partner: is the partner that collects the personal data from the data subjects and has the first and main contact with them. In personal data processing activities where more than one Partner is involved, apart from the Data Partner, all involved partners are defined



transcribed and then the recording medium erased or destroyed. No video recordings will be used as part of formal data collection. The appropriate institutional Data Protection officer (DPO) will be made aware of, and consulted if needed or required, if any issues related to research ethics arise during the course of the project.

The project contains a dedicated work package (WP8) to make sure the project complies with ethical principles and relevant regulations, on a national and on a European level, such as the GDPR. All personal data processed in the project will be secured adequately and kept confidential.

ECO-READY will appoint an Ethical Board (ETB) consisting of external independent experts. The Ethical Board's (ETB) main purpose will be to provide independent ethical expertise to ensure the project's activities respect the Do No Significant Harm (DNSH) principle as defined in the "EU Taxonomy" in Regulation (EU) 2020/852, including data collection and handling. The ETB will advise and assist the EB and ECO-READY partners in understanding and appropriately addressing the ethics issues raised by their activities, and to report independently to the Commission.

Artificial Intelligence: the AI models in the project will not process data sets containing personal data. There will be no direct interaction with the AI system by end users. Results of the AI systems are predictions of future food safety risks, which will support risk managers to take timely mitigating measures. Therefore, any negative impact on human rights and values is very limited, whereas the positive impact is improved food safety. The AI-powered ECO-READY components will explicitly exhibit the following qualities and behaviour with respect to the EU Ethical AI requirements, specifically on what relates to Data Management:

- **Privacy and Data governance:** Any data collected or generated through ECO-READY will have a license defined in the Data Management Plan. Appropriate documentation will be provided to ensure that data management is transparently defined for the users of ECO-READY.
- **Transparency:** Any AI process taking place in the data manipulation components of ECO-READY will be clearly indicated, along with any potential limitations/risks for employing the process.
- **Diversity, non-discrimination and fairness:** Any AI algorithm or process of ECO-READY will be performed on agnostic data, thus making sure that no gender or socioeconomic information is considered for the results of the algorithms.
- **Societal and environmental well-being:** ECO-READY is oriented towards social inclusion in environmental and other aspects. In terms of the developed AI, ECO-READY will employ data from the Living Labs, which will ensure the training of the algorithms with relevant data.
- **Accountability:** ECO-READY will keep traces of data that are stored or linked to the system and used for AI purposes, thus allowing the reproducibility of any decision made using AI to facilitate any future audit or explanation on how decisions are taken.

Consumers' surveys and personal data collection: All interviews and online surveys for consumers' data collection will include procedures for informed consent of the participants



and will be conducted in compliance and according to the EU General Data Protection Regulation. Moreover, personal data protection procedures will be defined according to the ECO-READY data management plan.

All participants of the Living Labs and project events will have the right to obtain the erasure of personal data relating to them and the abstention from further dissemination of such data according to the GDPR. They will be informed about this right in the information sheets. Applications for erasure of data will be carried out without delay. In case the personal data has been made public, the consortium will take all reasonable steps, including technical measures, to inform third parties which are processing such data, that a data subject requests them to erase any links to, or copy or replication of that personal data. A procedure for exercising the right to be forgotten and to erasure will be provided.

