REDONDO

REVERSIBLY DESIGNED CROSS-LINKED POLYMERS

D8.1
DATA MANAGEMENT PLAN

Work Package 8
[Communication, Dissemination and Exploitation]

Prepared by RDC Informatics

Deliverable Information

Responsible Partner:	RDC Informatics
Work Package:	8
Contributing Partner(s):	all
Dissemination Level:	PU - Public
Type:	R — Document, report
Due Date:	28.02.2023
Submission Date:	28.02.2023
Version:	1.0

Project Profile

Program	HORIZON EUROPE
Call	HORIZON-CL4-2021-RESILIENCE-01
Topic	A DIGITISED, RESOURCE-EFFICIENT AND RESILIENT INDUSTRY 2021 (HORIZON-CL4-2021-RESILIENCE-01)
Number	101058449
Acronym	REDONDO
Title	Reversibly Designed Cross- linked Polymers
Start Date	01/09/2022
Duration	48 Months

Document History

Version	Date	Author	Remarks
1	27.02.2023	Dionisis Koutsantonis Konstantinos Simopoulos	

EXECUTIVE SUMMARY

This deliverable (D8.1 Data Management Plan) aims to describe the processes followed in the REDONDO project regarding the development of a Data Management Plan. Under the requirements and guidelines of Horizon Europe, projects must focus from an early stage on how research data will be stored and the scientific outputs validated. These requirements foresee to:

- Develop (and keep up-to-date) a Data Management Plan (DMP).
- Deposit the data in a research data repository.
- Ensure third parties can freely access, mine, exploit, reproduce and disseminate this data.
- Provide related information and identify (or provide) the tools needed to use the raw data to validate research results.

Several types of datasets will be produced; hence data sharing and communication between the partners will be established via a DMP and a platform (Redondo OIE – Open Innovation Environment) designed specifically for this purpose. According to the Data Mapping scheme, data and metadata generated from disparate sources will be combined into meaningful and valuable information. Maximum capabilities of data sharing and metadata usage will be achieved, considering Data Security, Privacy, and General Data Protection Regulation compliance. This deliverable is developed according to Horizon Europe Data Management Template (version 1.0, 05 May 2021) provided by EU Grants and it will be updated continuously during the project duration.

Table of contents

LIST OF FIGURES

LIST OF TABLES

TABLE OF ABBREVIATIONS

Abbreviation	Definition
CA	Consortium Agreement
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Protection Regulation
HE	Horizon Europe
IP	Intellectual Property
LOD	Linked Open Data
OADR	Open Access Data Repository
OIE	Open Innovation Environment
RDF	Resource Description Framework

1. Introduction

1.1. About REDONDO

REDONDO (Reversibly Designed Cross-linked Polymers) is a Horizon Europe project (Project No: 101058449) that aims to produce recyclable, reversibly cross-linked polyethylene. Indeed, cross-linked polyethylene is a difficult to recycle material. REDONDO aims at achieving a fully reversible cross-linking process that will enable the production of sustainable-by-design cross-linked polyethylene and can further be applied to other cross-linked polymers.

1.2. REDONDO & Open Science Practices

The adoption of **Open Science Practices** in the European research and development sector has been one of the most solid and fundamental pillars on which EU-funded research projects operate. Many are the benefits to be considered, as not only the society can be well informed and (even) directly engaged in the research and innovation process but also the general public and in particular European citizens and taxpayers can be evidently assured of a transparent, well-directed and efficient EU R&D policy. From the researchers' point of view, open science practices have a tremendous impact on the ability of our society to innovate, removing barriers and rendering research results and technological advancement findable, accessible, interoperable and (re-)usable (FAIR principles). Throughout the past years, the adoption of open science practices has improved the quality, efficiency, transparency and social acceptance of research in Europe.

REDONDO consortium will adopt and promote Open Science Practices to the maximum extent. All open access material as well as scientific publications resulting from the project will be issued as open access, in compliance with the Horizon Europe guidelines, by utilizing the **Open Access Data Repository** (OADR) **Zenodo** (https://zenodo.org/), connected with **OpenAire** (https://www.openaire.eu/) and with the **Open Research Europe** platform (https://open-research-europe.ec.europa.eu/). The goal is to ensure the integrity and the reproducibility of the project results, as well as to facilitate any derivative or relevant works using, integrating and/or building on REDONDO results. One of the most important and highly impactful aspects of open science adoption in REDONDO is related to its Safe- and Sustainable- by design characteristics. SbD and SusbD assessments along with best-practices handbooks as well as a big part of REDONDO's post-project exploitation involve open science concepts as they are aiming for those technological advancements to be applicable by extension to other polymers and other end-use applications. Criteria, best practice handbooks and related deliverables will be published as openly as possible to support open science principles and facilitate its benefits ensuring no harm to the commercial exploitation of the project's results. This is one of the main pillars of the project itself, crucial for its implementation and a key part of its impact maximization measures.

1.3. REDONDO Data Management Plan

Managing the research data collected and processed during an innovative multi-disciplinary project such as REDONDO requires careful consideration. Different types of data will require unique data governance procedures, including technical and open science policy requirements for data collection, management, preservation and sharing. The **Data Management Plan** (DMP) is a key element that describes the data management policy that will be used by the consortium partners regarding the data collected, processed and/or generated within the project.

REDONDO Data Management Plan (DMP) is a formal document that contains all the necessary activities that will be pursued for efficient organization and data arrangement from different disciplines to provide accessibility to all data generated from partners and process modelling in REDONDO's testing facilities and thus hold valuable information by reducing costs and time. For REDONDO's datasets to meet the requirements to be findable, accessible, interoperable and reusable (FAIR), the DMP further adopts the essential task of Data and Metadata Handling. Metadata summarizes crucial information for identifying mechanisms, data cataloguing, approach to keywords, clear versioning, identifying sources, data curation, open availability, vocabularies, discovering relationships between data, and further regulating the information produced during the REDONDO project. Moreover, a sustainable DMP requires contemplating challenges in Data Security and Data Privacy for data to maintain high technical and scientific quality. Aspects such as storage, transfer, and recovery of sensitive data will be exploited within DMP. All the above-described activities are harmonized with the **General Data Protection Regulation** (GDPR) guideline.

The DMP clearly describes all the steps that consortium partners have to follow in order to achieve successful data management within the project and meet all goals for project results exploitation. In Figure 1, these processes are distinguished and presented.

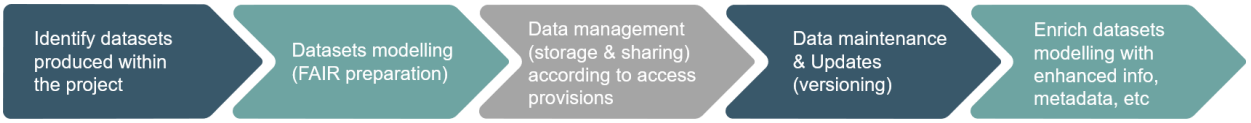

Figure 1. DMP provisions for successful data management within the REDONDO project

1. Identify datasets produced within the project
2. Model datasets to meet FAIR principles
3. Store and share actual digital data according to access rights provisions
4. Maintain and update data to the newest versions
5. Continuous optimisation of dataset modelling for optimum performance according to FAIR principles

In order to achieve efficient data management, RDC will provide under REDONDO Partners' Area in the website (available at https://www.redondo-project.eu/), a special section named **REDONDO OIE** (**Open Innovation Environment**) where digital tools for datasets modelling and data management will be available to partners. This document describes the use of these tools and will assist partners to achieve optimal data management processes as will be described in the next paragraphs.

2. Data Summary

2.1 Data collection/generation and its relation to the objectives of the project

REDONDO partners will create digital data as the project progresses. According to the CA of the project, a general overview of the data expected is presented in Table 1.

Table 1. General overview of REDONDO Data

WP	Processes	Type of data	Purpose
1	Project Management & Coordination	Reports (docs, presentations & spreadsheets)	Project Management & Reporting
2	Specifications and requirements	Numeric ranges	SbD, SusbD, rPEX and additive development
3	rPEX synthesis, characterization and reversibility	Numeric, Designs	SbD, SusbD, Replicability, Compounding
4	Additive formulation & integration	Numeric, Process	SbD, SusbD, Replicability, Compounding
5	Production, Validation	Numeric, Prototype, Process	SbD, SusbD, Certification, end-user exploitation, specs and requirements
6	Safety assessment	Numeric Specs	bD, Replicability, Best practices
7	Sustainability assessment	Numeric, Specs, Model, Process, Database	SusbD, Additives, Recyclability Replicability, Best practices
8	Communication, dissemination and exploitation	Reports (docs, graphics, video, presentations & spreadsheets)	Project dissemination & exploitation activities

In order to describe in detail all expected data from each partner within the project, every partner will fill before the 6M consortium meeting the questionnaire presented in Annex I "**REDONDO DATA RADAR – A DMP QUESTIONNAIRE**" and this section of the DMP will be updated accordingly in next version.

2.2 Data Modelling / Handling / Sharing

All digital data (files) of REDONDO will be categorized into datasets. The dataset is a critical ontology in

every DMP as its structure and definition are critical in order to achieve data management according to FAIR principles. The data handled in the project will be documented in a **Dataset Template** (see Annex II), which has been created to compile the most relevant information on the dataset lifecycle. The Dataset Template includes the following sections:

1. **Data summary**
2. **FAIR data**
    - **2.1. Findable data**
    - **2.2. Accessible data**
    - **2.3. Interoperable data**
    - **2.4. Reusable data**
3. **Allocation of resources**
4. **Data Security**
4. **Ethical and legal aspects**

In Annex III of DMP a **Dataset Index** of the datasets already identified will be created and will be updated accordingly. All Annexes will be updated with new datasets throughout the project duration.

RDC prepared a complete methodology to communicate any changes in data management as well as handle and share data, as presented in Figure 2.

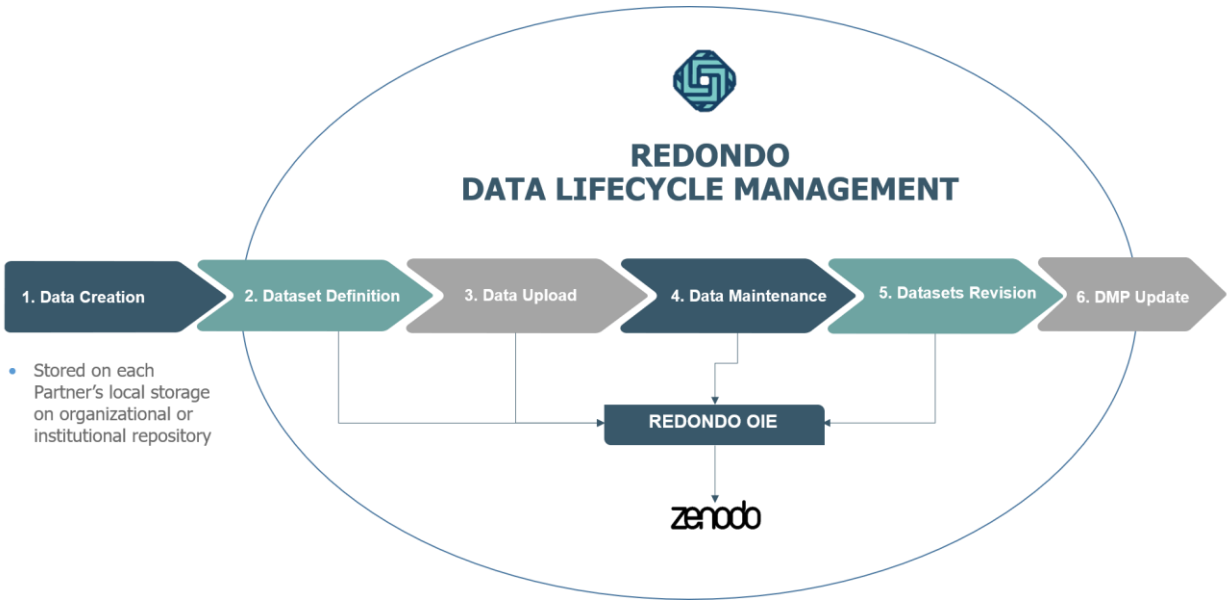

Figure 2. REDONDO DATA LIFECYCLE MANAGEMENT

The methodology consists of the following steps:

1. **Data Creation**
   Each partner creates digital data regarding the REDONDO project. These data are stored and manipulated according to organizational/institutional internal provisions regarding data management.

2. **Dataset Definition**
   In order to start the data lifecycle within REDONDO a partner has to define a dataset. Redondo

datasets are structured based on the dataset template presented in Annex II. These dataset templates can be declared online through the appropriate interface of REDONDO OIE.

3. **Data Upload**
   After defining dataset types, partners can upload, manage, distribute, and share data among partners and the public (if applicable) through REDONDO OIE. Each user has to declare the dataset type to which digital data belongs, as well as the data access rights among the following options:
   - Organizational/Institutional (data available to users of the same partner)
   - Among consortium partners (data available to users of all consortium partners)
   - Open Access (Project website)
   - Open Access Data Repository (Zenodo)

   REDONDO OIE will be integrated with Zenodo, so publishing data to Zenodo can be done automatically from each partner through the use of REDONDO OIE.

4. **Data Maintenance**
   The data upload mechanism through REDONDO OIE supports versioning for data, so it is easy to maintain data by uploading new versions of the same data.

5. **Datasets Revision**
   As the project progresses partners will be able to update the definition of the dataset through REDONDO OIE, enrich it with more metadata, define more searchable keywords, etc.

6. **DMP Update**
   As partners have the option to update an existing or create a new dataset anytime through REDONDO OIE, Annex III of DMP will be automatically updated by retrieving the datasets index from REDONDO OIE.

Thus, by utilizing REDONDO OIE and the processes described in this document all consortium partners will have a complete methodology to drive successful data management until the end of the project. This methodology and tools will be presented to partners during the 6M meeting and after finalizing the process REDONDO OIE will be the main data repository for REDONDO partners.

Through REDONDO OIE partner should have the option to publish directly on Zenodo under the Redondo Community operating at https://zenodo.org/communities/redondo/

3. Fair Data

Data management within the REDONDO project will be done according to FAIR data principles to make data Findable, Accessible, Interoperable and Reusable.

3.1. Making data findable, including provisions for metadata

Standard identification mechanisms

Digital Object Identifiers (DOI) will be used for each data upload on REDONDO OIE. These DOIs will be utilized for REDONDO OIE users only. All data that will be published on Zenodo will receive also a DOI from this repository for future reference.

Naming conventions

In order to provide a unique and persistent identification in REDONDO datasets, the following name convention will be used as nomenclature for dataset identifiers:

REDONDO_PARTNER_YYY_ver_ZZ

where:

- PARTNERX: Partner Name Acronym (e.g. AUTH, RDC,....)
- YYY: Dataset Number for each partner (e.g. 001, 002,....)
- ZZ: Version (e.g. 01, 02, 03,....)

Search keywords to optimize data re-use

All REDONDO datasets and file uploads will be enriched with search keywords. These keywords will be treated as metadata and they will able to make data more findable and reusable.

Clear version numbers

Declaring the version of a dataset or/and data file(s) in REDONDO is critical, so the REDONDO OIE is armed with a version management mechanism.

Metadata standards

Metadata summarize valuable information; hence it will be crucial in REDONDO's ecosystem for data cataloguing, identifying sources, discovering relationships between data and also organizing the information that will be produced during the project. The Resource Description Framework (RDF) is a metadata data model that enables the encoding, exchange and reuse of structured metadata. Linked Open Data (LOD) is a way of publishing structured data like RDF, that allows metadata to be connected and enriched, so that different representations of the same content can be found and links to be made more useful with the related resources through semantic queries.

3.2. Making data openly accessible

Data openly available as the default

Section 2b of the Dataset Template provides information related to the accessibility of data. For each dataset, the partner that owns the data will indicate whether the dataset will be made openly available.

If certain datasets cannot be shared (or need to be shared under restrictions), the partner will explain why, clearly separating legal and contractual reasons from voluntary restrictions. The openly accessible data (including the associated metadata and documentation) will be deposited in an open access data repository.

Restrictions for dataset sharing
The consortium agreement ensures the protection of the Intellectual Property of all partners. Data that are produced during the REDONDO project will be published upon the consortium's agreement whenever sensitive information is not included.

Specific beneficiaries to keep their data closed
The IPR arising from the results of the project will be the property of the partner that has developed it. Procedures and tools that will be produced will remain on their respective owners for exploitation.

How will the data be made accessible
All project data will be stored in REDONDO OIE and some of them in Zenodo. REDONDO OIE will operate on cloud infrastructure managed and maintained by RDC.

Methods and software tools to access the data
Data access can be seamless via methods such as an API, but the REDONDO OIE will provide a simplified search interface for easy access via data & metadata querying. The suggested metadata type in REDONDO will follow a simple text format (comma separated) accessed via elastic search technologies.

Data access committee
RDC is responsible for data governance during the REDONDO project duration. The Project Management Board will be responsible to define data management after project finalization, and policies defined hereinafter.

Conditions for access
Machine-readable licenses and access requirements (i.e. pdf files) will be available for platform users in REDONDO OIE.

Identification to access the data
User identification during access is one of the main tasks that the data security of the REDONDO project handles. Users will be granted access to the OIE platform by the consortium. In the Zenodo repository, sharing will be made possible only by the approval of the depositor of the original file.

3.3. Making data interoperable

Interoperability of data
As much as possible, the partners are encouraged to use non-proprietary, open and standard formats for various types of data widely used by the research community.

Data and metadata vocabularies, standards or methodologies to follow
Section 2c of the Dataset Template will indicate what data and metadata vocabularies, standards or methodologies will be followed to make the data interoperable.

3.4. Increasing data re-use

Data licensing
The last section related to FAIR data will indicate the conditions for data re-use, particularly after the end of the project. The partners will define the possible licensing terms and restrictions for each dataset, particularly for third-party use.

Timeframe for data access
During the project, the data shall be made accessible to the consortium partners under the terms of the Grant Agreement and Consortium Agreement. This section will also indicate the timeframe of data availability (when would be made available for re-use and for how long).

Use by third parties
Beneficiary data usage by third parties during and after the REDONDO will be ensured according to the current status of consortium agreements on data management.

4. Allocation of Resources & responsibilities

RDC will be responsible to provide the necessary tools, methodology and support to all partners to manage and share their data effectively during the project. RDC will be also responsible to download the index of the dataset from REDONDO OIE and update regularly the Annex III of this document for future versions. Nevertheless, each partner of the REDONDO project will be responsible for the safe storage of their own datasets. Each partner will be responsible for the identification of new datasets that will be handled within their own tasks. Each partner will be responsible for filling in the Dataset form in REDONDO OIE. If the dataset will be made openly accessible, the dataset owner will be responsible for uploading it to Zenodo (https://zenodo.org/) through REDONDO OIE. The consortium has agreed with the use of this free-of-charge repository for making the data accessible (without precluding the further use of other repositories, e.g., institutional). For the publications, the consortium will publish them in scientific journals that allow open access.

5. Data Security

Provisions for data security

Zenodo repository stores user passwords using strong cryptographic password hashing algorithms (currently PBKDF2+SHA512). Zenodo has a 12-hourly backup cycle with one backup sent to tape storage once a week, while a daily backup cycle will be set for cloud servers of RDC. According to the backup plan of RDC Infrastructure, 7 full infrastructure backups are maintained for the last 7 days, on a daily routine. Each consortium partner will also be responsible for the data security and the secure storage of their own datasets in their institutional repositories.

Long term preservation and curation

The digital datasets that will be used/generated within the project will be safely stored and managed by expert IT system administrators of each organization. As general guidelines, the data will be stored in at least two separate locations to avoid data loss; the partners will limit the use of USB flash drives; and the files will be labelled systematically in order to ensure consistency of the final dataset.

6. Ethics

No ethical or legal issues that can have an impact on data sharing have been identified for now. These issues will be monitored during the execution of the project and, if needed, taking into account in the next versions of the DMP.

7. Other Issues

At the moment, the REDONDO project does not make use of other national, funder, sectorial or departmental procedures for data management and no other direct issues have been found in this initial phase of the data management plan.

8. Conclusions

REDONDO's Data Management Plan will be continuously reviewed and updated to provide efficient and accessible data sharing and communication between the partners, throughout the projects's lifespan. RDC will provide an integrated Open Innovation Environment platform tailored to REDONDO's needs and specifications, accelerating project's development with data safety, undisrupted availability and a plethora of data collaboration tools and features.

Annex I – Redondo Data Radar Questionarie

REDONDO

REDONDO DATA RADAR - A DMP QUESTIONARIE

CONTINOUS REPORTING TOOL FOR DATA GENERATED DURING THE PROJECT

Partner Name:
Abreviation:
Completed by:
Contact e-mail:
Date:
Vesion: 1.0 (6M)

DATA SUMMARY								
Relation to the objectives of the project:	**WP1**	**WP2**	**WP3**	**WP4**	**WP5**	**WP6**	**WP7**	**WP8**
What is the type of data? E.g. Text files, Numerical - Multimedia, Models, Software								
What is the format of data? (xlsx, doc, pdf, jpg, etc.)								
What is the expected size of the data that you intend to generate or re-use? (please define in MBs/GBs etc)								
What is the method of capture of data? (e.g. experimental data, market search, surveys…)								
Will you re-use any existing data (through clarifying licences)?								
Any ethical (related to human, personal data, animal) issues can related to this data?								
Might data be useful ('data utility'), outside the project?								
Define Data Access Rights	Organizational/Institutional	Organizational/Institutional	Organizational/Institutional	Organizational/Institutional	Organizational/Institutional	Organizational/Institutional	Organizational/Institutional	Organizational/Institutional
DATA REPOSITORIES								
Does your organization uses a Data Repository?	NO							
If the answer above is YES, please define the link to re posítory								
Do you have any objection to use Zenodo Repository as Open Access Repository for Redondo project	NO							
ETHICAL & LEGAL ASPECTS								
Please indicate if you have identified any ethical or legal aspect that could impact the data sharing.								

Annex II – Redondo Dataset Template

REDONDO DATASET DEFINITION FORM

Dataset Owner(s):	
Creation date:	
Version:	
Dataset Identifier:	
	Use the convetion name **REDONDO_PARTNER_YYY_ver_ZZ** Where: PARTNERX: Partner Name Acronym (e.g. AUTH, RDC,....) YYY: Dataset Number for each partner (e.g. 001, 002,....) ZZ: Version (e.g. 01, 02, 03,....)

1. DATA SUMMARY

Dataset description

Describe the dataset in a few lines.

Purpose of the data / Relation with the project

Explain the purpose of the data collection/generation and its relation to the objectives of the project.

Type of data regarding the source

☐ Observational: data captured in real-time, typically outside the laboratory, e.g., sensor readings, telemetry, survey results, images

☐ Experimental: data collected under controlled conditions, e.g., gene sequences, chromatograms, magnetic field readings

☐ Simulation: machine-generated data from test models, e.g., climate models, economic models

☐ Derived/Compiled: generated from existing datasets, e.g., text/data mining, compiled database, 3D models

☐ Reference or canonical: static or organic collection (peer-reviewed) datasets, most probably published and/or curated, e.g., gene sequence databanks, chemical structures, census data, spatial data portals

Type of data regarding the form

Select the way the information is provided and generated.

☐ Text, e.g., field or laboratory notes, survey responses

☐ Numeric, e.g., tables, counts, measurements

☐ Audiovisual, e.g., images, sound recordings, video

☐ Models, computer code

☐ Discipline-specific, e.g., FITS in astronomy, CIF in chemistry (please specify):

☐ Instrument-specific, e.g., equipment outputs (please specify):

Format of the data

Select the format in which the data are released.

☐ Text: plain text (TXT), HTML, XML, PDF/A

☐ Databases: XML, CSV

☐ Image: JPEG, JPG-2000, PNG, TIFF

☐ Audio: AIFF, WAVE

☐ Containers: TAR, GZIP, ZIP

☐ Other (please specify):

Origin of the data

If the data is re-used, please explain how.

☐ Newly collected/generated

☐ Re-used (please explain):

Expected size of data

Give an estimate, e.g., a few kB/MB/GB, etc.

_____

☐ Fixed: never change after being collected or generated

☐ Growing: new data may be added but the old data is never changed or deleted

☐ Revisable: new data may be added, and old data may be changed or deleted

Utility of the data

Indicate to whom the data might be useful. E.g., partners, stakeholders, citizens, etc.

2a. FAIR DATA: MAKING DATA FINDABLE

Standard identification mechanism and metadata

Are the data discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g., persistent and unique identifiers such as Digital Object Identifiers)? What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Naming convention

Indicate any naming convention that you follow.

Keywords

Provide keywords to optimize possibilities for re-use.

2b. FAIR DATA: MAKING DATA OPENLY ACCESSIBLE

Dataset accessibility

If confidential, explain why, clearly separating legal and contractual reasons from voluntary restrictions.

☐ **Confidential data** ☐ **Public** data

Reason for confidentiality: ***click here to enter text***

Repository to be used to deposit the dataset

What repository will you use to make the data accessible, including deposition of the data and associated metadata, documentation and code. Note that all open data must be deposited at least in Zenodo.

☐ **Zenodo**

☐ **Institutional (please specify):**

☐ **Other (please specify):**

☐ **None**

Methods or software tools needed to access the data

Indicate what software is needed to access the data.

2c. FAIR DATA: MAKING DATA INTEROPERABLE

Explain what data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable.

2d. FAIR DATA: INCREASING DATA RE-USE

If applicable, define the data licensing approach to permit the widest re-use possible. Indicate the chosen license tools.

If applicable, define the timeframe for making data available for re-use. Indicate if any embargo period is required.

Indicate any restrictions of data re-use by third parties, particularly after the end of the project.

Indicate for how long it is intended that the data remains re-usable.

Explain how data quality is assured.

3. ALLOCATION OF RESOURCES

If applicable, indicate any allocation of resources to data management.

4. DATA SECURITY

Indicate the provisions in place to ensure data security.

5. ETHICAL & LEGAL ASPECTS

Please indicate if you have identified any ethical or legal aspect that could impact the data sharing.

Annex III – Redondo Dataset Index

Dataset ID	Owner(s)