



RETROFIT SOLUTIONS TO ACHIEVE 55% GHG REDUCTION BY 2030

Data Management Plan

WP 9 – Project Management
Task 4 – Data Management Plan
D9.2 – Data Management Plan
Partners involved: CNR, all
Authors: Emanuele Spinosa, Cecilia Leotardi, Stephanie Elisabeth Seidl, Alessandro Iafrati





Project details

Project Title	RETROFIT SOLUTIONS TO ACHIEVE 55% GHG REDUCTION BY 2030
Project Type	Innovation Action
Project Acronym	RETROFIT55
Grant Agreement No.	101096068
Duration	36 M
Project Start Date	01/01/2023

Deliverable information

Status (F: final; D: draft; RD: revised draft)	F 30/06/2023
Planned delivery date	30/06/2023
Actual delivery date	30/06/2023
Dissemination level: <ul style="list-style-type: none">• 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	PU
Type: Report, Website, Other, Ethics	Report



Document history

Version	Date	Created/Amended by	Changes
01	12/06/2023	Emanuele Spinosa, Cecilia Leotardi, Alessandro lafrati (CNR)	First draft
02	28/06/2023	Reuben Dsouza, Elissavet Klonopoulou (SFWD), Thomas Koch (AES), Stephanie Elisabeth Seidl (CNR)	Additional details on file numbering, repository and metadata formats
Final	30/06/2023	Alessandro lafrati (CNR)	Final version

Quality check review

Reviewer (s)	Main changes / Actions
Reuben Dsouza, Elissavet Klonopoulou (SFWD)	Suggestions on how to manage the responsibilities for the data management and on the data-sharing platform connected to the project webpage Specifications
Stephanie Elisabeth Seidl (CNR)	Suggestions on the trusted repository to be chosen and on the format of data/metadata. Contribution on the types of licenses available for the publications and on the link between availability of data and publication date
Alessandro lafrati (CNR)	Final check



Table of Contents

Executive Summary	5
List of Acronyms	6
1 Introduction	7
2 Data Management in the framework of RETROFIT55 Project	8
2.1.1 Objectives of the RETROFIT55 Data Management Plan	8
2.1.2 Partners role and responsibilities in implementing the DMP	8
2.1.3 Data Management Plan updates	9
2.1.4 Data Management Plan and General Data Protection Regulation	9
3 Data Summary	10
4 FAIR data	11
4.1 Making data findable and inter-operable, including provisions for metadata	11
4.1.1 Input data files	11
4.1.2 Output data files	13
4.2 Making data Accessible	15
4.2.1 Repository	15
4.2.2 Data Sharing Platform	16
4.2.3 Confidential/Public Data	16
4.2.4 Data Availability Statement	17
4.2.5 Data availability policy	19
4.3 Data Re-use policy	20
5 Other research outputs	21
5.1 Publications	21
5.2 Web based tool	21
5.3 Presentations, Posters etc.	21
6 Data security	23
7 Ethics	24
8 Other issues	25
Closing remarks	26
References	27



Executive Summary

This Data Management Plan (DMP) describes the data that will be collected and generated during the RETROFIT55 project. The main focus is on how the data will be managed, from the acquisition to the exploitation and dissemination, as well as on the way in which the data will be made available for future use. This document will be updated during the project lifetime: this first release presents an overview of the data that are expected to be collected/generated during the project and the way the Consortium intends to manage them. The report will also provide information about accessibility and re-use of other project outputs, such as publications, the project web-page, presentations, and other dissemination material. The next versions will provide more details about the data management and how it is implemented in the project, as well as a detailed description of all datasets collected and generated during the project

Since RETROFIT55 involves either industrial data or the development to higher TRL of innovative industrial solutions, which may eventually turn into market products, some data used in the project and some of the results will not be open to the public or they will be made available in a suitable non-dimensional form, which will still allow to quantify the benefits of the developed solutions but will prevent the possibilities to reverse engineering.



List of Acronyms

CA	Consortium Agreement
DMP	Data Management Plan
DOA	Description Of the Action (Annex 1 of the GA)
DOI	Digital Object Identifier
EC	European Commission
EU	European Union
FAIR	Findable Accessible Interoperable Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
GHG	Green-House Gas
ICT	Information and Communication Technologies
IPR	Intellectual Property Rights
KPI	Key Performance Index
PC	Project Coordinator
PO	Project Officer
WP	Work Package
XML	eXtensible Markup Language



1 Introduction

RETROFIT55 is a Horizon Europe Innovation Action aimed at delivering new technologies and solutions for the reduction of the fuel oil consumption and Green-House Gas (GHG) emissions. These solutions will be implemented on existing ships by retrofitting. In addition, Information and Communication Technologies (ICT) strategies will be developed to support the ship owners during the retrofitting process.

In order to achieve these objectives, the use of wings/sails for Wind Assisted Ship Propulsion (WASP) and the injection of air bubbles at the bottom of the hull to reduce the skin friction through an Air Lubrication System (ALS), will be investigated. These solutions will be complemented by an optimization of the route, a hydrodynamic optimization of the hull, the introduction of green energy solutions and of a smart energy management module. Besides the technical solutions, a web based configuration tool will be developed with the purpose of guiding and assisting the ship owners and ship builders to configure the retrofit interventions, by combining different technologies. The different configurations will be evaluated in a life-cycle perspective in terms of specific Key Performance Indexes (KPI). To this aim, surrogate models will be used to represent the different alternatives in a digital form.

Design and technical data, together with data acquired at a high sampling rate on board of existing ships will be made available for different ship types by the ship owners participating to the project. Such data will be used to define specific use cases and will represent an incredibly valuable tool for the development and validation of the digital twins of the different subsystems, as well as for the quantification of the gain achievable when installing new energy saving solutions and green technologies. Besides the real data, other data and surrogate models will be generated during the project as a result of numerical simulations, laboratory measurements and, possibly, on-board measurements. The benefit of the implementation of the new technologies will be determined either numerically or via laboratory and large scale experiments.

In order to fulfil the EU requirements, the definition of a number of principles and a clear organization of the data management are necessary and will be summarized in this report, i.e. the Data Management Plan (DMP). The DPM is outlined at the beginning of the project (month 6) and will be then updated regularly until the end of the project, when the final version will be delivered.

This report is organized as follows: in Section 2 the objectives of the Data Management Plan are described, as well as the responsibilities of the project partners in terms of data management. In Section 3 the types of data collected and generated during the project are summarized. In Section 4 the provisions to make data Findable, Accessible, Inter-operable and Re-usable (FAIR) are outlined. In Section 5 the other research outputs of the projects will be illustrated, which are mainly scientific publications and presentations for dissemination events. In Section 6 provisions for data security are outlined, in Section 7 ethical questions are briefly recalled and, finally, in Section 8 other issues are addressed.

2 Data Management in the framework of RETROFIT55 Project

2.1.1 Objectives of the RETROFIT55 Data Management Plan

The objective of the DMP is to provide indications for the project partners in terms of data collection, generation, organisation and storage, in order to make the data available and usable by the Consortium and by potential users outside the project, in compliance with the confidentiality and the Intellectual Property Right (IPR) policy described in the Consortium Agreement and at the dissemination level indicated in the Grant Agreement.

Specifically, the DMP provides the guidelines that will ease the coherent and consistent full life-cycle data handling throughout project lifetime, according to privacy and ethical issues and complying with the guidelines of the European Commission (EC), i.e. in line with the FAIR protocol.

The indications in the present report concern the different stages of data management and include information on what kind of data are collected, generated, and on the way, they are handled during and after the end of the project, also addressing storage and preservation tasks.



Figure 1: Key phases of data management.

Specifically, the DMP includes the following specific indications on the key phases of the data management (see Figure 1).

- Collection and generation of data: origin, nature and characteristics (type, content and size) of the collected/generated data;
- Storage: how and where data will be stored and who will be responsible for security, maintenance and access;
- Access & Sharing: format and data structure, how data are made available and shared within the Consortium and outside the project, and who determines the level of sensitivity of the data to be shared;
- Archiving: how data are made available after the end of the project.

2.1.2 Partners role and responsibilities in implementing the DMP

All project partners are involved in data management in one way or another. However, some of them have specific roles and responsibilities:

- the Project Coordinator (*i.e.*, CNR), has the responsibility to ensure that data are stored in a trusted repository (in this case Zenodo.org) and to guarantee the respect of the FAIR principles, assuring that the proprietary data are secured;
- each WP leader is responsible for the integrity and the compatibility of the data generated within the WP they are responsible of;
- each Task leader is responsible for the quality of the data generated in the task and in the registration of the metadata.

2.1.3 Data Management Plan updates

The DMP is a “living” document, which will be updated continuously during the project, and new versions will be formally delivered at M18 and M36. The next versions will take into account updates in terms of the data collected and generated during the project and in terms of the data management policy (e.g. new data, changes in consortium policies, changes in consortium composition, changes in applicable EU regulation, IPR aspects, etc.).

2.1.4 Data Management Plan and General Data Protection Regulation

This DMP is developed in line with the General Data Protection Regulation (EU) 2016/679 (GDPR) which was adopted on the 14th of April 2016, and became enforceable as of the 25th of May 2018. Another reference document is the [Guidelines 4/2019 on Article 25 Data Protection by Design and by Default](#), published in October 2020.

3 Data Summary

As mentioned in Section 1, different data types will be used during the project and generated as project's outcomes. Data are preliminarily classified into **input data** and **output data**.

- **Input data (collected data)**, are provided either by
 - the **ship owners** involved in the project (e.g., Laskaridis, Grimaldi) and can be delivered, for example, in terms of ship design data (characteristics of the ships, propulsion system, manoeuvring and stabilization devices, etc.) and of on board energy systems data (e.g. main engine, diesel generators, energy distribution, energy absorbers, etc.). Moreover, operational data in terms of time histories acquired at the highest possible frequency, will be provided in order to tune the digital twins of the different use cases. Example are: ship speed and route, weather conditions, sea state, wind speed and direction, fuel oil consumption, energy absorption of the auxiliaries, GHG emission, etc.
 - the **ship repair and conversion** company (Astander) in terms of data to support the life cycle assessment of the different retrofit interventions;
 - **other partners** in terms of additional information for the specific use cases, for the various technologies considered in the project and for the design and characterization of the retrofiting;
 - **literature** (publications, papers, reports) or **public databases** (e.g. data about the Japanese Bulk Carrier that will be employed for the testing).
- **Output data (generated data)** will be delivered as a result of the project activities in terms of numerical simulations and analyses, laboratory experiments, data assimilation, data driven and surrogate models. Possible examples:
 - global values, e.g. ship resistance, propulsion thrust and torque, total and averaged power absorption, sails loads, ALS efficiency, averaged power absorption, etc.;
 - local data, e.g. hydro-aero dynamics fields, green energy production, energy storage, air fraction in ALS, optimized routing, etc.
 - KPIs characterizing the different combination of the retrofiting solutions evaluated through surrogate models for the selected use cases. These data will be made available to the public via demo-cases in the **web-based platform**.

A detailed list of the types of data generated in the project will be provided at the next update of the DMP at M18 and in the final version due by M36.

4 FAIR data

4.1 Making data findable and inter-operable, including provisions for metadata

In order to make the data files easily findable and inter-operable (*i.e.* integrated with other data and with applications or workflows for analysis, storage, and processing) they will be grouped into **datasets**. All data files in each dataset must have the same format and must be homogeneous, *i.e.* they all must be read/interpreted using the same procedure/algorithm. Each data file in the dataset will be assigned to a unique **ID code**. In addition, each dataset will be associated with a unique **metadata file**, also characterized by a unique ID code. The metadata file will provide information on the creator, date of publication, the types of data files and, more importantly, all the necessary information on how to read the data files in the dataset. The metadata will be compliant with the DataCite Metadata standard scheme, which is the standard also adopted by Zenodo.org.

As mentioned in Section 3, the data files can be distinguished into two categories: output data files and input data files. The file nomenclature is different according to this classification.

4.1.1 Input data files

Input or collected data files are provided by the project partners or come from literature or external sources, but they are not generated during the RETROFIT55 project. These data files are associated with the project partner that collects or provides the data. For example, they could be related to the ship design parameters, geometry, propeller shape, on-board systems, operational data of a ship on several routes, etc.. The association between the owner of the input data and data/metadata files is shown, as an example, in Figure 2. It is anticipated here that some of the data could be industrially relevant and will not be disclosed due to IPR issues.

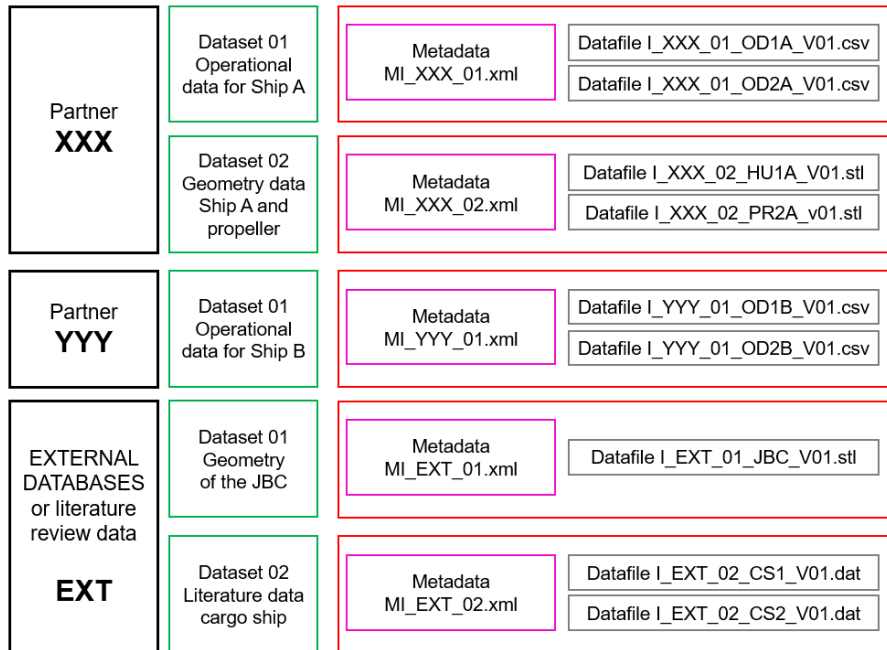


Figure 2: Example of association of data providers to metadata/data input data files

These data files are associated with an ID code in the following format:

I_[P]_[K]_[ABC...]_V[JJ].ext

where

- [I] is the letter which identifies the file as an INPUT data file
- [P] is the partner ID code, according to Table 1. If the data come from an external source, literature review or from other databases, the ID is set as EXT
- [K] is the dataset identifier (two-digit)
- [ABC...] is an optional free text to identify in short the type of data. Details on the content are provided in the corresponding metadata file
- V[JJ] is a running two-digit index that indicates the version of the data file, in case of revisions/corrections (e.g. V01, V02 etc.). The highest number indicates the latest version
- .ext is the extension of the file.

Table 1: Partners ID

Organisation name	Partner ID
Consiglio Nazionale delle Ricerche	CNR
Aalto University	AALTO
Atlantec Enterprise Solutions	AES
Friendship Systems	FSYS
National Technical University of Athens	NTUA
Astilleros De Santander	ATD
SimFWD	SFWD
Rina Services	RINA

Laskaridis Shipping	LASK
Bound4Blue	B4B
Liverpool John Moores University (Associated)	LJMU
Armada Technologies Ltd (Associated)	ARM
Advanced Wing Systems	AWS
Grimaldi Euromed	GRI
External source (details in metafile)	EXT

The list of input datasets will be included in the next versions of the DMP.

Each input data file will be associated with a single **(input) metadata file**, which is a sort of readme file, in XML format, containing all the necessary information related to the data files. The metadata file will be compliant with the DataCite standard scheme, which will make it easily findable (following the policy of Zenodo.org). A template of the XML metadata file, editable with any text editor, will be prepared and made available on the Microsoft Teams repository. The input metadata file will also be associated with a unique identifier in this format:

MI_[P]_[K].xml

where

- MI is a code that identifies the file as a metadata of the INPUT data files (dataset)
- [P] is the partner ID as listed in Table 1
- [K] is the dataset identifier (two-digit).

In this way, **the association between the data file and the metadata file** in the corresponding dataset is easily recognizable from the part **[P]_[K]** of the filename. For instance, looking at Figure 2, data files I_XXX_01_OD1A_V01.csv and I_XXX_01_OD2A_V01.csv are associated to the same metadata file MI_XXX_01.xml and vice-versa.

The metadata file must provide the information on what type of data files are associated with the digit [K] as well as the information about the software needed to open them (e.g. Excel, Matlab, Tecplot, Simulink, Paraview, etc.). The metadata file must also include complete details about the data content. For instance, if the data refer to operational data, information about the ship type, route, date, list and order of the variables, and any other parameter needed to identify unambiguously the data and to extract them from the file must be provided. Each metadata will also contain a few **keywords**, which will make the metadata file and the related data files easier to find, especially by possible third-party users.

4.1.2 Output data files

The **output data files** are generated within the project, hence they are associated with the work packages and the tasks in which they are produced. Figure 3 shows an example of a possible nomenclature for the output data for a given WP.

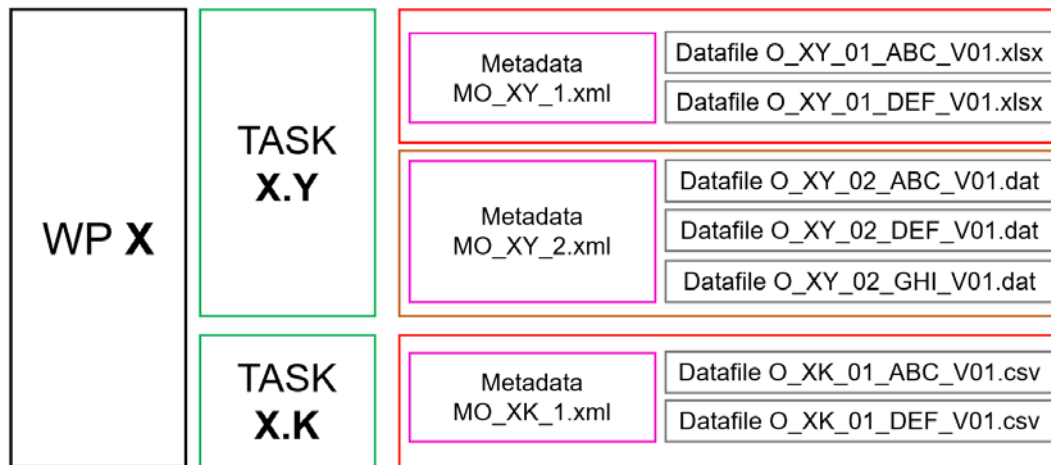


Figure 3: Example of association between WPs, tasks, metadata and output data files

Each data file is associated and named using a code with the following format

O_[WT]_[K]_[ABC...]_V[JJ].ext

where

- O is the letter which identifies the file as an OUTPUT data file
- [WT] is a two-digit number indicating work package (1 to 9) and related task
- [K] is the dataset identifier (two-digit)
- [ABC...] is an optional free text to identify in short the type of data. Further details on the content are provided in the corresponding metadata file
- V[JJ] is again a running two-digit index that indicates the version of the data file (e.g. V01, V02 etc.). The highest number indicates the latest version
- .ext is the extension of the file.

The list of datasets pertaining to all WP and tasks will be included in the next versions of the DMP.

Each data file will be associated with a single **(output) metadata file**, which has the same XML format and characteristics of the input metadata files. The output metadata file is associated with a unique identifier in this format:

MO_[WT]_[K].xml

where

- MO is a code that identifies the file as a metadata of the OUTPUT data files (dataset)
- [WT] is a two-digit number identifying work package and task
- [K] is dataset identifier (two-digit)
- .xlsx is the format of the metadata file.

In this case, **the association between the data file and the metadata file** in the corresponding dataset is easily recognizable from the digits **[WT]_[K]** of the ID code. For instance (see Figure 3),

the metadata file MO_X1_01.xml is associated with the data files O_XY_01_ABC_V01.xlsx and files O_XY_02_DEF_V01.xlsx.

The output metadata file must provide the information on what type of data are provided in the associated data files concerning origin of the data (e.g. experiments, numerical simulations, etc.), the software used to generate them, format, variables, the software to be used to open the files and any other useful information, e.g. use case, simulation conditions, keywords, etc..

It is worth noticing that, as for the input files, not all files can be made open to the public as some of them include industrially sensitive information which cannot be disclosed to the public. In those cases, whenever possible, data will be provided in dimensionless form.

4.2 Making data Accessible

4.2.1 Repository

The data and metadata files will be stored in the repository **ZENODO.org**. A community named RETROFIT55 on the Zenodo platform will be created, which is a space on the website that will group all the files related to the project. A Digital Object Identifier (DOI) for each dataset or any submitted file will be automatically assigned by the platform. When the datasets or any other type of document is submitted to Zenodo, a metadata file, compliant with the Horizon Europe requirements, will be automatically produced. This is required by the regulations of the Horizon Europe programme to ensure the transparency of the data according to the FAIR principles.

Different types of data can be uploaded on ZENODO platform, namely:

- publication
- poster
- presentation
- dataset
- image
- video
- software
- lesson
- other

Zenodo accepts up to 50GB per dataset, multiple datasets uploads are permitted. There is no size limit for communities. As an example, Figure 4 shows a screenshot of the webpage of a Zenodo community related to a Horizon 2020 project.

Each task leader or partner who produces/provides the data must upload the single datasets following the nomenclature specified in Section 4.1, so that he can be considered the primary holder of the Intellectual Property Rights. All the necessary information to build the metadata file, required by Zenodo.org, must be provided during the submission process. The Project Coordinator (CNR) will supervise and help the authors (WP/ task leaders or other project participants) to submit the data appropriately and by the due time.

It is worth noting that once a file is uploaded on ZENODO platform, **it cannot be easily removed or modified.**

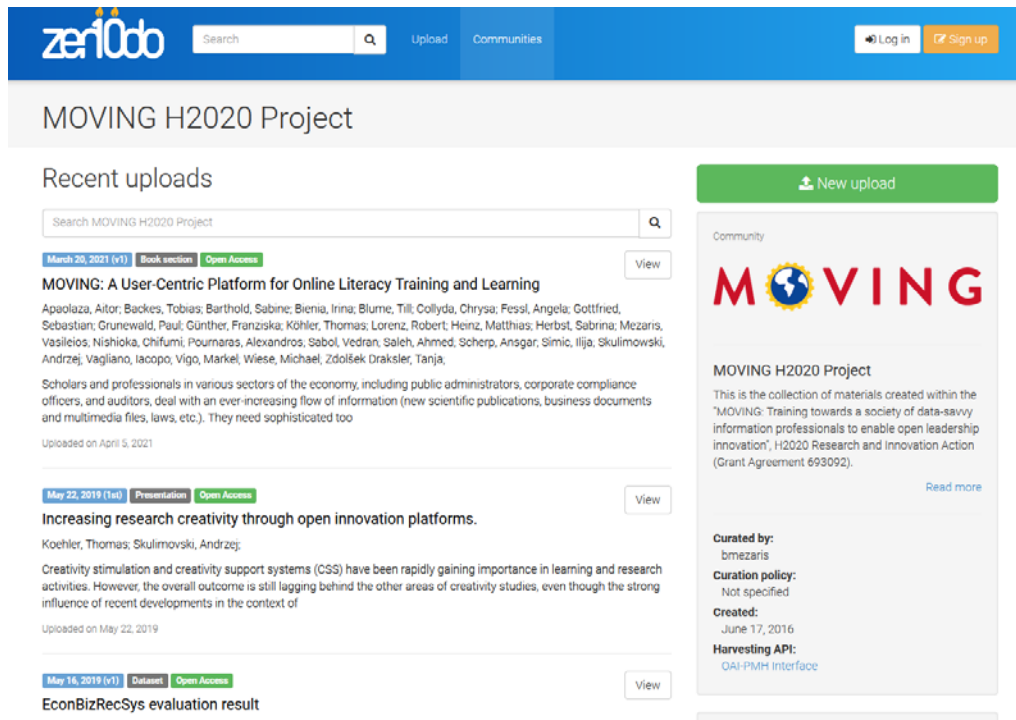


Figure 4: Screenshot of the web-page of a community on the Zenodo platform, related to a H2020 project (<https://zenodo.org/communities/moving-h2020/>).

4.2.2 Data Sharing Platform

A backup of all datasets will be also present on a data-sharing platform, only accessible by the project partners, inked to the project website Members Area. The platform will contain all the datasets, which can be shared among the project partners. The data stored in this platform will not be accessible by external users.

4.2.3 Confidential/Public Data

The data files collected and generated during the RETROFIT55 project are classified according to three level of confidentiality, as specified in the GA. These levels are the following:

- **SC** → strictly confidential: these are company sensitive data, provided that there is no legal basis for being distributed or diffused;
- **CC** → currently confidential: these data are confidential during the project development, but they can possibly become public at a later stage e.g., Intellectual Property Rights (IPR) relevant data;
- **P** → public: this type of data can be distributed at any time and also re-used by third part users (external to the project) upon the publication of the project results.

The type of confidentiality level must be specified for each dataset. This will be detailed in the next versions of the DMP.

According to the regulations of Horizon Europe, to each non-confidential dataset or document, (i.e. CC or P data) a Creative Common 4.0 licences, type CC0 (Public Domain Dedication) or CC-BY has

to be applied. A Creative Commons (CC) license is a copyright license that enable the free distribution of a copyrighted "work", i.e. any creative material made by a person (a painting, a graphic, a book, a song and its lyrics, a photograph etc.). A CC license is applied when the author wants to provide other people with the right to share, use, and build upon a created work, see for more information the website <https://creativecommons.org/licenses/?lang=en>. A CC license allows a certain degree of flexibility to the author, who can choose, for example, to allow only non-commercial uses of his work. In addition, a CC license protects the people who use/redistribute the author's work from the concerns of copyright infringement, as long as they abide by the conditions specified in the license itself. In particular, as indicated in the CC website:

- the **CC0 license** “enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law”;
- the **CC-BY license** “lets others distribute, remix, adapt, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials”.

Other types of CC licenses are not in line with directives of Horizon Europe and cannot be chosen.

4.2.4 Data Availability Statement

A data availability statement, which refers to any used generated or collected datasets, must be added to the end of any article or document (e.g., papers, posters, presentations, etc.) before the reference list section. The statement must describe each dataset, include a link to the repository containing the data, and the DOIs.

The statement must specify if a CC0 Public Domain Dedication or a CC-BY 4.0 license is applied to document and to the related datasets.

There can be exceptions in case data cannot be shared because of legitimate interest, such as industrial exploitation, confidentiality, trade secrets, security rules, union competitive interests or IPR, including patents. In this case, the article/document must include a description of the restrictions on the data and all necessary information required for a reader or reviewer to apply for access to the data and the conditions under which access will be granted.

A few examples of Data Availability Statements, which is adapted from the website <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines>, is presented in Table 2.

Table 2 - Example of Data Availability Statement

DATA TYPE	DATA AVAILABILITY STATEMENT EXAMPLE	DATA CITATION EXAMPLE
Data deposited into a generalist repository	Open Science Framework: Japan PVD 2018. https://doi.org/10.17605/OSF.IO/QW2AE ; Registration DOI https://doi.org/10.17605/OSF.IO/7Y4AV (Yamada, 2020) This project contains the following	Buljan I, Marusic A, Pina D. Ethics issues identified by applicants and ethics experts in Horizon 2020 grant proposals: Supplementary files. Open Science Framework. 2021. https://osf.io/cyp4w/ <i>Example taken from: Buljan I, Pina DG and Marušić A. Ethics issues identified by applicants and ethics experts in Horizon 2020 grant</i>

	<p>underlying data:</p> <ul style="list-style-type: none"> • PVDJapan2018.xlsx. (The dataset.) • Description of the Dataset.txt <p>Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).</p>	<p><i>proposals. F1000Research 2021, 10:471</i> (https://doi.org/10.12688/f1000research.52965.2/)</p>
Data deposited into a repository with accession codes	<p>The underlying data has been deposited in the ProteomeXchange Consortium via the PRIDE partner repository, accession number PXD027611: https://identifiers.org/pride.project:PXDO27611/</p>	<p>Wright, J and Choudhary, J. Identifying and characterizing Thrap3, Bclaf1 and Erh direct interactions using cross-linking mass spectrometry. PRIDE. 2021. https://identifiers.org/pride.project:PXDO27611/</p> <p><i>Example taken from: Shcherbakova L, Pardo M, Roumeliotis T and Choudhary J. Identifying and characterising Thrap3, Bclaf1 and Erh interactions using cross-linking mass spectrometry. Wellcome Open Res 2021, 6:260</i> (https://doi.org/10.12688/wellcomeopenres.17160.1/)</p>
Data with access restrictions	<p>LSHTM Data Compass: Treatment of child wasting: Child Health Research Initiative (CHNRI) prioritisation exercise dataset, https://doi.org/10.17037/DATA.00001882</p> <p>This project contains the following underlying data:</p> <ul style="list-style-type: none"> • Underlying data file 1: dataset (NWL-CHNRI-dataset) (restricted access) • Underlying data file 2: dataset description (NWL-CHNRI-dataset-codebook) (unrestricted access) <p>Due to the fact that open posting of data on a repository was not included in the study information sheet at the time the survey was done, data access will be granted once users have consented to the data sharing agreement and provided written plans and justification for what is proposed with the data. Data access may be obtained by submitting a request to the No Wasted Lives, Action Against Hunger authors via the LSHTM Data Compass repository. Requests will be reviewed by Action Against Hunger/ No Wasted Lives (the lead agency for this study) and key collaborators as named on the repository.</p>	<p>Kerac M, Angood C, Mayberry A, <i>et al.</i>: Treatment of child wasting: Child Health Research Initiative (CHNRI) prioritisation exercise dataset. LSHTM Data Compass. 2020. http://www.doi.org/10.17037/DATA.00001882</p> <p><i>Example taken from: Angood C, Kerac M, Black R et al. Treatment of child wasting: results of a child health and nutrition research initiative (CHNRI) prioritisation exercise. F1000Research 2021, 10:126</i> (https://doi.org/10.12688/f1000research.46544.1)</p>
Articles without data	No data associated with this article	None required



Articles where the data consists of bibliographic references	The data for this article consists of bibliographic references, which are included in the References section.	Standard bibliographic references
--	---	-----------------------------------

4.2.5 Data availability policy

The RETROFIT55 Consortium supports the idea of Open Science, and the benefits coming from the data re-use at large scale. Although some of the technologies and solutions are protected by IPR and cannot be fully disclosed, most of the data generated in the project will be made openly available, in dimensionless form where appropriate. The policy for the open access of data is summarized in the schematic in Figure 5. If there is no confidentiality issue or legitimate interest, the datasets must be uploaded on an open access repository, which is Zenodo.org for this project. If the data are underpinning a publication, they can be uploaded **at the latest at the time at the publication**, otherwise they must be uploaded as soon as possible, but in any case **before the end of the project**. In cases of public emergency, the granting authority can ask for immediate open access of the data. If exceptions to open access to data apply, data should be made available at least to the legal entities that need the data to address the public emergency. Only in exceptional cases, data can be deposited after the project has ended. For more information, see the website <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm>.



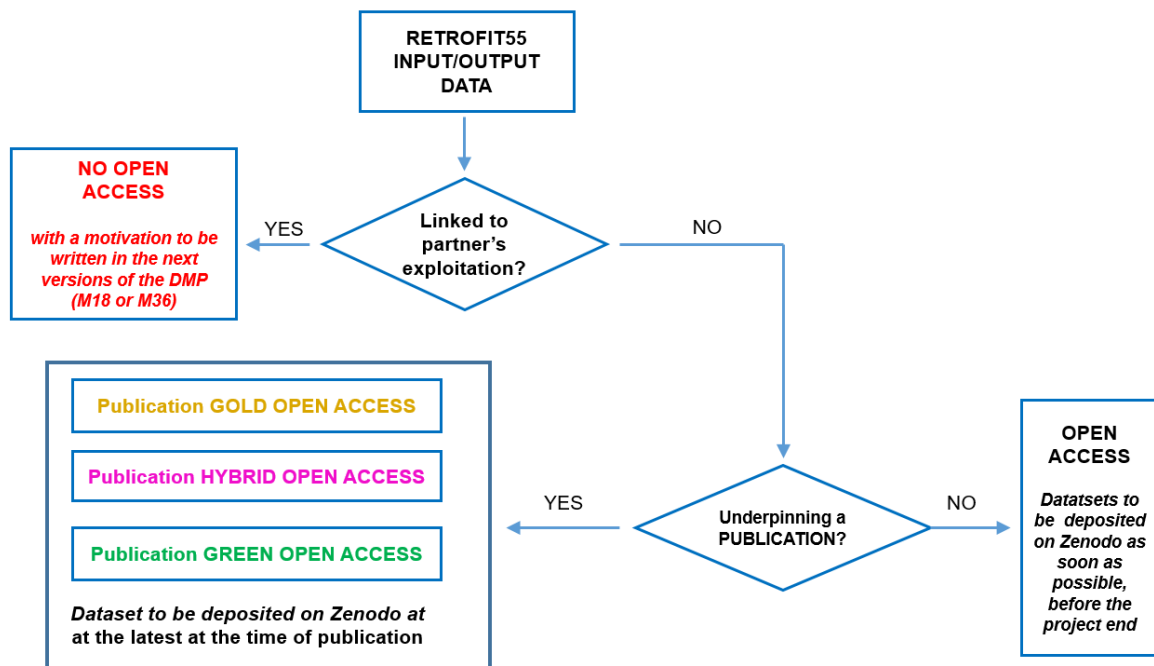


Figure 5: Policy for the open access of the data generated in the project RETROFIT55

4.3 Data Re-use policy

In the framework of Horizon Europe, a clear documentation must be provided to validate the data analysis and facilitate the data re-use. To this purpose, the metadata files produced for each dataset must contain all the information to read and interpret the data files, not only for the project partners, but also for possible external users.

The data which are not classified as confidential will not be shared and therefore cannot be re-used by users. The public data are reusable and can be downloaded from the Open Access repository Zenodo.org, as of the date specified in Section 4.2.5.

5 Other research outputs

5.1 Publications

In the RETROFIT55 project, scientific publications are among the main research outputs. The procedure for the publication acceptance is detailed in the GA and summarised in the Data Management Handbook, Deliverable 9.1.

As specified in the DOA-GA, all scientific publications must follow the Open Access policy. For more information, see the web-page <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-publications>.

When choosing a journal to publish the RETROFIT55 results, following the policy of Horizon Europe, three options are available:

- **Gold Open Access**, in which the final published version of your article is made permanently and freely available online for anyone, anywhere to read/download. Gold Open access allows the author to share their research anywhere as of the publication date, while keeping the author's copyright. An Article Publishing Charge (APC) is typically required by the journal, which is eligible for reimbursement.
- **Hybrid Open Access**, in which the article can be published on a submission journal, which also offers the option for open access after the payment of an APC. In this case the APC is not eligible for reimbursement in the framework of Horizon Europe, therefore this option should be chosen preferably if the institution/company has already a read-and-publish deal with the selected editor/journal.
- **Green Open Access**, also known as self-archiving, is the procedure when the author posts on a public repository an earlier version of the manuscript sent to a submission journal, i.e. the accepted version after the peer reviewing process. This enables the author to share the article, complying with the EU regulations, without having to pay any APC. In Horizon Europe **no embargo period is permitted**, therefore it is necessary to choose a journal/editor complying with this requirement.

Each paper must contain a data availability statement and a link to the underlying datasets, as specified in Section 4.2.4. If it is not possible to share the data, for the reasons mentioned above, the data availability statement must report this.

Just like the datasets, the manuscript shared in the public repository must be associated with a Creative Commons CC0 or CC-BY license (<https://creativecommons.org/licenses/?lang=en>). It is worth to be remarked again that other types of licenses, such as CC-BY-NC or CC-BY-ND or CC-BY-NC-ND are not in line with the policy of Horizon Europe and cannot be chosen.

5.2 Web based tool

The web-based tool will be available to the public with some specific examples developed within the project. The background information or other data that are liable for commercial exploitation and that could limit the patentability of the developed technology, will not be made accessible to the public.

5.3 Presentations, Posters etc.

The main results of the project will also be presented at events, conferences or dedicated meetings with stakeholders.





6 Data security

As far as data security is concerned - including data recovery, as well as secure storage/archiving and transfer of sensitive data - Zenodo is hosted at CERN, which is an Intergovernmental Organization (IGO), with its seat in Switzerland and therefore not subject to the EU's General Data Protection Regulation (GDPR). Data protection at CERN, which Zenodo complies with, is governed by CERN's Operational Circular 11 (OC11) that offers data protection at the same high standards and comparable to GDPR.

The security system of the data-sharing platform will be in line with the directives of Horizon Europe.

The task leaders, the work package leaders must also perform regular local data backups to minimize at most the risk of data loss.



7 Ethics

At the current stage there are no ethics or legal issues that can have an impact on data sharing. Any possible issue will be discussed in further meeting, whether any problem arises.



8 Other issues

At the current stage, the project is not supposed to make use of other national, funder, sectorial, departmental procedures for data management.



Closing remarks

In this report, the preliminary structure of the Data Management Plan of the RETROFIT55 project is provided. Some general information concerning the type of data used and generated by the project are provided, along with the organisation of the file repository (Zenodo.org), the way used to name the files and the associated metadata files, and the approach to the Open Access. The document will be updated during the project, and a final version is due by the end of the project, in which all the details concerning the data produced in the project and the way they can be accessed for reuse will be included.

References

1. RETROFIT55 Grant Agreement.
2. RETROFIT55 Consortium Agreement.
3. WP9 Kick-off meeting presentation
4. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf
5. <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines>
6. <https://creativecommons.org/licenses/?lang=en>
7. <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm>
8. Guidelines 4/2019 on Article 25 Data Protection by Design and by Default