



D1.4 Deliverable Data Management Plan

Enzymatic CO₂ capture
in a rotating packed bed
and electrocatalytic CO₂
reduction to useful
products



Reuseproject.eu



Funded by
the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101172954. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union.

Document control sheet

The project	REUSE: Enzymatic CO ₂ capture in a rotating packed bed and electrocatalytic CO ₂ reduction to useful products
Grant Agreement n°	101172954
Call identifier	HORIZON-CL5-2024-D3-01-05 - Development of carbon fixation technologies for biogenic flue gases
Type of action	HORIZON-RIA
Coordinator	CERTH
Project start date	1 October 2024
Project duration	36 months
Work package n° and title	1 Project management and coordination
Work package leader	CERTH
Dissemination level	PU - Public
Lead Beneficiary	CERTH
Author	Panagiotis Kazepidis
Reviewer(s)	All partners
Review date	26-03-2025
Due Date	M6

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

Index

1. Introduction	5
2. Legal provisions regarding data and IP in the CA.....	5
3. Terms used to describe data	6
4. Data processing.....	6
5. FAIR principles and questionnaire	7
6. Contributions from partners to questionnaire	8
7. Open access provisions and sharing during the project.....	20
8. Provisions for data from surveys	22
9. Ethics and Privacy	22
10. Partners	23

Executive Summary

In this deliverable we are presenting the data management plan (DMP). This DMP complies with the “Guidelines on FAIR Data Management¹”. This report provides specific information related to data management requirements that will be applied in the project. It pertains to issues regarding how the data will be gathered, maintained and shared by the partners following the FAIR principles. Key features of the DMP are the following:

- The data will be processed in a 3-stage procedure that includes gathering and organization, quality assessment and long-term archiving.
- A detailed questionnaire is answered by all partners that includes several questions under the categories of data naming, data description, standards and metadata, data sharing, data archiving and preservation. The key conclusion is that the generated data will be shared under the consortium agreement rules, that some generated data will be published based on golden or green access models and that the files are not of significant volume and can be easily maintained in the IT infrastructures of the corresponding organizations indefinitely.
- Specific provisions have been taken for data that will be generated from surveys described as part of the project activities, such data will abide by General Data Protection Regulation (GDPR)² rules among other provisions.

¹ <https://www.ffg.at/en/europe/heu/legal-financial/open-access>

² Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

1. Introduction

The aim of the Data Management Plan (DMP) is to define the type and the descriptions of all the data sets collected and generated by the project. This will cover how the research data will be handled during the project and how they will be stored, curated and accessed after the project. All research activities carried out in REUSE will be conducted in compliance with fundamental ethical principles. The project consortium and all partners will fully comply with:

- The Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regards to the processing of personal data and on the free movement of such data, and until valid, the repealing Directive 95/46/EC (General Data Protection Regulation),
- Directive 2002/58/EC on privacy and electronic communications,
- The Charter of Fundamental Rights of the EU (2000/c 364/01).

The DMP describes in detail all the data that will be generated and preserved, during the REUSE project, following specific methodology and standards, whether and how the data will be exploited or be made accessible for future use. Furthermore, the DMP outlines which methodologies and standards will be used in the data creation and management. All hosted data will comply with GDPR. We will also explain how and when the data will be shared and made open for re-use. In case parts of the research data cannot be openly shared, the DMP will provide proper motivations and data availability statements. The Plan will also be consistent with the CA, as well as the IPR and exploitation policy.

2. Legal provisions regarding data and IP in the CA

The consortium Agreement (CA) has very clear terms in section 8 about who owns and how the project results are handled. As the results are expressed mainly by data, this is a very relevant section. In particular, section 8 of the CA provides information on who owns the results, how they can be shared and jointly owned, how they can be transferred and how they can be disseminated, among other provisions and details. Section 9 is also very relevant as it determines the access rights. Access rights refer to the background, the granting of access rights for implementation and for exploitation, the access rights to entities under the same control and to partners entering the consortium, as well as specific provisions for access rights to software. Section 10 of the CA is also very relevant as it refers to how information can be disclosed from disclosing parties to recipients. The CA has been signed by all parties before the start of the project.

3. Terms used to describe data

There are several general terms that may be used to describe data. They are presented here for clarity:

Research data: This includes data that have been collected, observed, generated, created or obtained directly as a result of the project activities and/or from third-party sources. The research data may be the direct output of an activity or an intermediate output that is processed and analyzed to produce final research results.

Open research data: Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated, free of charge for our users. Such data will mainly refer to openly available deliverables, based on the Grant Agreement (GA).

Secondary data: These are data that pre-exist.

Open access data: Open access is understood as the principle that data should be accessible to relevant users, on equal terms, and at the lowest possible cost. Access should be easy, user-friendly and, if possible, Internet-based.

Metadata: Metadata are data used to describe other data. It summarises basic information about data, which can make finding and working with instances of data easier.

Research data repositories: Research data repositories are online archives where research data are stored and accessed. They can be subject-based/thematic, institutional or centralised.

4. Data processing

When generated, the data will be processed by their creators, prior to resulting in the final data set that could be stored and/or shared. This includes 3 main stages:

Stage 1: This stage pertains to the generation and/or gathering of data and their initial placement in a repository. At this stage, raw data should be organised systematically, using appropriate naming and dating, with sufficient descriptions of the circumstances, methods, instruments, software etc. used for their generation or gathering.

Stage 2: This stage involves the quality assessment of the data. Quality assessment pertains to the evaluation of the raw data and the decision as to whether they should be preserved for subsequent analysis and use.

Stage 3: This stage pertains to the preparation of the data for long-term archiving, e.g., for publications or patents. In this stage, it is important to analyse and clarify who is responsible for the data, what the data contain, what has been done with the data, who can use them and for which purpose. These metadata will help determine data ownership and their subsequent use, based on the provisions of the CA. When the data is transferred to a data repository, these are expected to be accompanied by relevant metadata that are

usually assigned by the host of the repository. For example, for publications that are accompanied by supplementary information or data files, the publisher automatically assigns the metadata that are necessary.

5. FAIR principles and questionnaire

The final version of the DMP will be submitted in M6, as contractually agreed in the GA. In this respect, the partners are asked to fill in the questionnaire shown in Table 1. This questionnaire summarizes all the information that is necessary to highlight the plans of each partner for the generated data and how they can be used.

Table 1: Questionnaire that should be completed by the partners.

1. Data Set and Reference Name
1.1 Identifier for the data set produced.
2. Data Set Description
2.1 How was data generated or collected? Which is the baseline data – if available?
2.2 What is the type/format and the scale of this data?
2.3 Did similar data already exist? If yes, is there the possibility of integration and reuse (if available)?
2.4 What kind of problem can this data set solve? What are the business/technology benefits for the stakeholders?
2.5 Has this data already been used in publications? If yes mention details.
3. Standards and Metadata
3.1 How can this data be found and used in an interoperable manner?
3.2 Which kind of metadata are you using (if any)?
4. Data Sharing
4.1 How will this data be shared?
4.2 Which are the main data access procedures?
4.3 Which are the technical mechanisms for dissemination and the relevant software to access /reuse the data?
4.4 Which are the general accessibility features? Will this data be public or private?
5. Data Archiving and Preservation (including storage and backup)
5.1 Which procedures will be put in place for data long-term preservation?
5.2 How long should data be preserved and what is its approximated end-volume?

6. Contributions from partners to questionnaire

Based on Table 1 above, Table 2 includes the contribution of the relevant partners in the specific questions. Partners that have activities involving data generation include CERTH, UAveiro, ULEIC, UNEW, NZ, CES, MMU, YSD, ETA and TBWR.

Table 2: Contributions from partners

1. Data Set and Reference Name
<p>1.1 Identifier for the data set produced.</p> <p>CERTH: 1. REUSE_Equilibrium_and_Kinetic_Data, 2. Pilot_Plant_Testing_Data, 3. Gas Diffusion Electrode and Catalyst Coated Membranes development and characterization data, 4. Membrane Electrode Assemblies development, characterization and testing (in CO₂R cell) data, 5. CO₂R prototype stack assembly and testing data.</p> <p>UAveiro: 1. Biomass: proximate and ultimate analysis, kinetic decomposition. 2. Gas outlet composition, Low heating Value, Carbon conversion, and gas yield (from gasification experiments). 3 Empirical data (same as obtained in 2) from statistical models.</p> <p>ULEIC: 1. Catalyst synthesis and characterization data for tar abatement. 2. Catalyst synthesis, characterization and testing data (from CO₂R experiments) 3. Corrosion assessment data (obtained in 2) for tin (Sn)-based catalysts 4. Empirical data (based on DOEs for 1, 2).</p> <p>UNEW: (1) Data on the design and construction of the composite rotating packed beds using CA doped materials. (2) Packing material data (active surface area, pressure drop). (3) CO₂ capture rate of CA-doped packing materials in three different solvents (KOH, AMP, DMCA).</p> <p>NZ: 1. Enzyme design, selection, and characterization; 2. Characterize new immobilization strategies for carbonic anhydrase (CA); 3. Prototype development and reusability assessment.</p> <p>CES: 1. Data gathering from pilot project and 2. Economic and technology assessments</p> <p>MMU: 1. Catalyst activation and synthesis data for CO₂R experiments. 2. Empirical data</p> <p>YSD: 1. Equipment dimensions and energy needs 2. Economic data 3. LCA data in scale-up studies.</p> <p>ETA: The data sets produced by the communication and dissemination activities of the project will include social media content, reports, factsheets, website publications, and event materials. Each data set will be assigned with a unique identifier, such as REUSE_Press Release 1_011124_v3 to ensure traceability and proper versioning of disseminated materials.</p> <p>TBWR: 1. Data gathering from pilot project and 2. Economic and technology assessments</p>

2. Data Set Description

2.1 How was data generated or collected? Which is the baseline data – if available?

CERTH: The data in data set 1 will be generated from equilibrium and kinetic measurements of solvent solutions that contain CA. The data set 2 will be generated from pilot plant tests. The baseline data for data sets 1 and 2 pertain to pure aqueous solvent solutions without the use of the materials developed in the project. The data in data set 3 and 4 will be generated from the results obtained in CO₂ reduction cell experiments. The data in data set 5 will be generated from the results obtained from the CO₂R prototype stack testing which is part of the pilot plant testing.

UAveiro: The data in data set 1 will be obtained from biomass characterization methods (ultimate and proximate analysis, thermogravimetric experiments). The data set 2 will be collected from gasification runs (3kWth and 75kWth gasifiers). The numerical data set is obtained using the experimental data from 2 and the statistical software Design Expert.

ULEIC: The data in dataset 1 will be obtained from experiments conducted using the catalysts developed for tar abatement in a gasifier. The data in dataset 2 will be generated from the results of FA formation rate and Faradic efficiency obtained via testing the developed tin-based catalysts in CO₂ reduction environments. The data in dataset 3 will be generated from the corrosion assessment of the developed tin-based catalysts in CO₂ environments containing phase-changing solvents.

UNEW: Data set 1 will be collected from documents used in the design and manufacture of the CA doped composite rotating packed beds. The data set 2 will be generated from pilot plant tests using CO₂ and NaOH. The baseline data will be data we collected previously using conventional packing materials. The data set 3 will be generated from pilot plant tests using different solvents. The decrease of CO₂ in the gas phase will be measured using CO₂ sensors. The increase of CO₂ in the liquid phase will be measured using titration method. The baseline data will be packing materials not doped with CA.

NZ: The data in data set 1 will be generated based on Novozymes library and enzyme engineering along with activity and stability assays. The data set 2 will be generated from lab-scale immobilization tests of the enzyme selected in data set 1. The experimental results obtained in data set 1 and 2 will be used to design a prototype (data set 3). Aqueous solvents will be used as baseline for data set 1 and 2 for the enzyme activity and stability.

CES: Data gathering and estimations considering data set 1 obtained during plant testing i.e., KPIs, CAPEX, OPEX and ongoing elaboration of data set 2 resulting in data like IRR, EIRR, NPV, levelised costs of units etc.

MMU: Modelling/simulation data will be generated using boundary conditions from combustion, corrosion and CO₂ conversion processes.

YSD: Data set 1 will be derived from tasks 6.2 and 6.3, where the different models that represent all the processes that take place in the REUSE system will be integrated and later from the control framework that will be proposed. Data set 2 will be the outcome of the techno-economic analysis of the optimal design of the REUSE system. One final data set, data set 3, will be generated from the LCA of the system.

ETA: The data is generated and collected through multiple communication and dissemination activities, including: 1) Updates from project partners on technology development and progress; 2) Content from project events, such as presentations, posters, recordings and participation feedback; 3) Reports, factsheets, and publications developed to communicate project outcomes; 4) Social media engagement metrics, website analytics, and stakeholder interactions.

There is no baseline data for communication and dissemination work within the scope of REUSE because this is a unique project with unique activities, development and results, thus all data will be generated as primary data.

TBWR: Data gathering and estimations considering data set 1 obtained during plant testing i.e., KPIs, CAPEX, OPEX and ongoing elaboration of data set 2 resulting in data like IRR, EIRR, NPV, levelised costs of units etc.

2.2 What is the type/format and the scale of this data?

CERTH: The generated data for all data sets are numerical values written in excel files. Their size is a few KB.

UAveiro: The generated data for all data sets are numerical values written in excel files. Their size is a few KB.

ULEIC: The numerical data generated will be recorded in Excel files. Graphs will be exported in text files (.txt, .csv). Correlation and categorical datasets will be provided in docx/xls format and image format (TIFF/JPEG). Numerical simulation data will produce files in .mph format. When proprietary software is used for the generation of data, the output data will be readily available in formats such as .csv, .txt, and .jpeg. Depending on the format, the data files could end up being a few MB.

UNEW: Data set (1) will consist of documents and drawings illustrating the cloth and packing designs. Their size will be up to 30 MB. The generated data for data sets (2) and (3) are numerical values written in excel files. Their size will be up to 20 MB.

NZ: The generated data for all data sets are numerical values, typically excel files ranging from KB to MB in size.

CES: Collated data are mainly in written figures supported by descriptions and comments.

MMU: XY data, contours (image file), csv files. When proprietary software is used for the generation of data, the output data will be readily available in formats such as .csv, .txt, and .jpeg. Depending on the format, the data files could end up being a few MB.

YSD: The generated data for all data sets are numerical values written in excel files. Their size is a few KB.

ETA: The data includes: 1) Text-based materials: reports, factsheets, blog articles, and social media posts (pdf, docx, html); 2) Visual content: Infographics, slides, and promotional materials (jpeg, png, pptx); 3) Multimedia: Video recordings of events, webinars, and interviews (mp4); 4) Analytical data: social media engagement metrics, website traffic statistics (csv, xlsx, json). The scale of data depends on the actual engagement level, and it will grow over time as more updates, events, and materials are produced.

TBWR: Collated data are mainly in written figures supported by descriptions and comments.

2.3 Did similar data already exist? If yes, is there the possibility of integration and reuse (if available)?

CERTH: As noted in point 2.1, previous data exist only for pure aqueous solvent solutions. Such data will be integrated with the new ones, as we will provide comparative diagrams containing both data types, to highlight performance differences with the developments of the project.

UAveiro: Data change as a function of the used biomasses. Even the same kind of biomass delivers different absolute values. Still, some trends are expected. Data for blended biomasses are scarce.

ULEIC: Previous data exist only for different catalyst combinations and will be integrated with the developed ones if they are comparable. Data on the catalysts for tar abatement and CO₂ reduction through the proposed routes is still scarce.

UNEW: (1) No existing data on the geometry of the cotton-based packing materials for RPB. (2) No existing data on the surface area and pressure drop of cloth based composite packing materials. (3) No existing data of the CO₂ capture rate of the CA-doped packing materials in RPB.

NZ: Engineered enzymes are proprietary of Novozymes and as such no data on either free or immobilized is currently available outside the organization.

CES: Data partly exist from previous project(s) but not including the present to be newly developed CCU (using new solvents, new designed RPB unit with new packaging material, CO₂R Cell etc.)

MMU: No simulation/modelling data.

YSD: Similar data exist for different solvents only for the absorber, but yet the generated data will be different for existing ones due to the novel approach of the project. However, the existing ones can be integrated with the new, for informative comparative analysis.

ETA: While there are similar projects funded under the same Call, this project is unique in its objectives, methodologies, technological developments, and activities to be organized. Therefore, the

communication and dissemination data must be generated based on the real progress of this specific project.

Existing data from similar EU-funded initiatives may provide useful background information, but they cannot replace primary data collected during the project. Instead, previous research and communication strategies may be referenced to enhance dissemination efforts, ensuring alignment with broader EU goals and maximizing the project's impact.

TBWR: Data partly exist from previous project(s) but not including the present to be newly developed CCU (using new solvents, new designed RPB unit with new packaging material, CO₂R Cell etc.)

2.4 What kind of problem can this data set solve? What are the business/technology benefits for the stakeholders?

CERTH: Data set 1 will include equilibrium behavior information for immobilized Carbonic Anhydrase (CA) in combination with amines that act as solvents in CO₂ capture processes. Such data are needed to assess the suitability of CA-amine solvent systems for CO₂ capture and to optimize CO₂ capture process design. Data set 2 will include pilot plant performance information with the materials developed in the project. The data sets 3 and 4 will provide insights into the optimum materials (catalysts, membranes, ionomers, gas diffusion layers, etc) and way of developing Membrane Electrode Assemblies (CCMs or GDEs) for CO₂R cells for the production of formic acid. Performance data of the CO₂R cell towards enhanced selectivity of formic acid and faradaic efficiency will be also produced. Data set 5 includes information on the stackability of the CO₂R cell and performance assessment of the prototype CO₂R cell. CERTH could provide such data from all data sets to equipment and process designing and manufacturing stakeholders.

UAveiro: One of the most important problems in thermochemical processes is the scattered data along the literature. Here, in REUSE, we deliver a systematic experimental procedure while comparing single and blended biomasses. Moreover, we deliver experimental data for scaling up by comparing 3kwth and 75 kWth gasifiers at similar operating conditions. Data are usually obtained using one-factor-at-time experiments. In REUSE, we use Design of Experiments to understand the interactions between key parameters and to target optimal performance and robustness.

ULEIC: The datasets (1, 2, 3) will provide insight into the screening of the catalyst combinations with optimum compositions that exhibit improved tar abatement and CO₂ reduction (when tested in an electrochemical cell). They will also provide test performance data of the CO₂R catalysts with enhanced selectivity towards formic acid and faradaic efficiency. Data set 3 includes information on the corrosion assessment and long-term performance of the tin-based catalysts in CO₂ environments. The benefits include the development of a scalable catalyst synthesis process for enhanced tar abatement in a gasifier, CO₂R to formic acid (with good selectivity) (from beaker scale to pilot plant scale).

UNEW: Data set 1 contains information on designs that integrate flexible cloth into a rotating packed bed. Rotating packed beds conventionally use rigid metal or plastic packing materials, which cannot be used for immobilising biological materials. Cotton cloth is a good substrate for immobilised enzyme but is not rigid enough to withstand rotational forces. By combining the cloth and metal packing, a composite rotating packed bed can be constructed that allows immobilised enzymes to be applied for carbon capture in compact process equipment. Data sets 2 and 3 will be the first attempt of cotton-based packing materials for RPB. The data set 3 will prove the feasibility of integrating enzyme in a RPB system. The integration of enzyme will greatly improve the kinetic of the reaction. It offers a possibility of running those reactions that conventionally have a slow reaction kinetic efficiently in RPB.

NZ: Data set 1 will include activity and stability information of CA necessary to select appropriate candidate(s) suitable for the immobilization and amine-based reaction. Building upon the findings of data set 1, the immobilization of CA will be optimized in data set 2 to offer good enzyme retention and residual activity on the selected supporting materials. Finally, enzyme reusability will be assessed for the final prototype providing critical information on the stability and longevity of the process.

CES: Establishing dedicated Observatory factsheets for D6.3:and Exploitation Plans for D7.2.

MMU: The modelled data set will complement experimental data work from combustion, corrosion and CO₂ conversion. Further insights into the combustion processes of biomass in the context of CO₂ formation and other pollutants will be obtained. CO₂ conversion route using sensitive analyses will be achieved.

YSD: Data set 1 is of great significance as it combines all the models that represent the different tasks of the system, allowing the simulation and optimization of the total system. Stakeholders can benefit massively by having the ability to estimate the equipment needed for industrial applications. This set is directly connected to the 2nd data set where the economic performance of the process will be estimated and then compared to the conventional approach on an economic level. Data set 3 is significant as it will provide information about the environmental impact and the viability of the system.

ETA: This data set helps address the challenge of effectively communicating and disseminating complex scientific and technological advancements to a broader audience, including policymakers, industry stakeholders, researchers, and the general public. Benefits for stakeholders include: 1) Increased awareness of project outcomes and technological advancements; 2) Facilitated knowledge transfer between research institutions, industry partners, and policymakers; 3) Enhanced stakeholder engagement for potential collaboration and exploitation opportunities; 4) Support for decision-making in policy and industry by providing accessible and well-structured data.

TBWR: Establishing dedicated Observatory factsheets for D6.3:and Exploitation Plans for D7.2.

2.5 Has this data already been used in publications? If yes mention details.

CERTH: This data set has not been used in publications.

UAveiro: This data set has not been used in publications.

ULEIC: This data set has not been used in publications.

UNEW: This data set has not been used in publications.

NZ: This data set has not been used in publications.

CES: This data set has not been used in publications.

MMU: No it has not.

YSD: This data set has not been used in publications.

ETA: As of now, communication and dissemination-related data of REUSE has not been used in major external publications. As the project progresses, it is expected to be referenced in the project reports, policy briefs, academic publications, and EU-funded research repositories. However, some documents have been published such as the 1st press release announcing the project kickoff, website, and social media post promoting the project launch and website launch.

TBWR: This data set has not been used in publications.

3. Standards and Metadata

3.1 How can this data be found and used in an interoperable manner?

CERTH: The data is confidential. If some of them are eventually published, they will be easy to find as they will be reported in tables or diagrams that will be included in the published papers. The data will be interoperable as they will be numerical values that could be adopted and used by any interested user.

UAveiro: Data will be easy to find as they will be reported in tables or diagrams that will be included in the published papers. The data will be interoperable as they will be numerical values that could be adopted and used by any interested user.

ULEIC: Data will be reported in tables, diagrams or images in the published papers. The data will be interoperable as they will be numerical values that could be adopted and used by any interested user.

UNEW: Data related to the performance of the CA is confidential. An agreement has been made between UNEW, Novonesis and CERTH, regarding the publication of the data. Before that, the data can be used internally for project development. Data on the performance of the cotton-based packing materials will be published subject to the agreement of commercial partners in the project. It will be interoperable as the idea can be easily adopted and used by any interested user.

NZ: The data is confidential. If some data are selected for publication purposes, it will be shared in the forms of tables or diagrams that will be included in the published papers. The data will be interoperable as they will be numerical values that could be adopted and used by any interested user.

CES: Meta-categories will be derived based on T5.4, T6.3 and T7.2 activities and eventually published according to the FAIR principles.

MMU: Data reported in standard formats (tables, graphs and images) in published papers and wider dissemination activity. The data will be interoperable as they will be numerical values that could be adopted and used by any interested user.

YSD: These data sets will be confidential. Some of them will be published and reported in tables and diagrams and discussed thoroughly in the text of the published papers. Since the data will be numerical values and could be adopted and used by any interested user.

ETA: The data generated through communication and dissemination activities will be made accessible via the project website and LinkedIn account to ensure findability and interoperability.

TBWR: The data is confidential. If some of them are eventually published, they will be easy to find as they will be reported in tables or diagrams that will be included in the published papers. The data will be interoperable as they will be numerical values that could be adopted and used by any interested user.

3.2 Which kind of metadata are you using (if any)?

CERTH: No metadata will be used. The data will be contained in the pdf or electronic versions of the printed, published material(s).

UAveiro: No metadata will be used. The data will be contained in the pdf or electronic versions of the printed, published material(s).

ULEIC: No metadata will be used. The data will be in the PDF or electronic versions of the printed, published material(s).

UNEW: No metadata will be used. All data sets will be generated in the laboratory and contained in pdf or electronic versions of the printed, published material(s).

NZ: No metadata will be used. The data will be contained in the pdf or electronic versions of the printed, published material(s).

CES: For the Exploitation Plan of T7.2 intention is to identify list of aggregated Key Exploitable Results (KERs) developed within the REUSE project.

MMU: No metadata will be used. The data will be in the PDF or electronic versions of the printed, published material(s).

YSD: No metadata will be used. The data will be contained in the pdf or electronic versions of the printed, published material(s).

ETA: To enhance the discoverability and proper documentation of the data, the following data will be applied: title, date of creation, authors/ contributors, versioning, keywords, data type and format, and access conditions.

TBWR: Meta-categories will be derived based on T5.4, T6.3 and T7.2 activities and eventually published according to the FAIR principles.

4. Data Sharing

4.1 What are the data accessibility features? Are they public or confidential / private?

CERTH: The data are currently confidential/ private. At the moment they can be accessed by their creators and by interested project partners under the provisions and terms of the grant and the consortium agreements.

UAveiro: The data will be public after publishing or when available in specific deliverables (with open access).

ULEIC: Suitable data will be available via publications (open access journals), including a cloud-based repository (Figshare). Data will also be available in specific deliverables with open access.

UNEW: The data are currently confidential/ private. At the moment they can be accessed by their creators and by interested project partners under the provisions and terms of the grant and the consortium agreements.

NZ: The data are currently confidential/ private and can be accessed by their creators and by interested project partners when deemed essential under the provisions and terms of the grant and the consortium agreements.

CES: All surveys and Workshop inquiries are openly circulated and responded of external stakeholders on voluntary basis as well as among the consortium in line with the work plan of DoA.

MMU: The data will be public after publishing or when available in specific deliverables (with open access).

YSD: The data are currently confidential/ private. At the moment they can be accessed by their creators and by interested project partners under the provisions and terms of the grant and the consortium agreements.

ETA: The accessibility of the data depends on its type. Materials with public dissemination level will be openly accessible via the website, social media and in the future relevant EU platforms. Meanwhile, restricted and confidential data such as partner's information, proprietary technologies, or unpublished research will remain confidential and will not be shared publicly.

TBWR: All surveys and Workshop inquiries are openly circulated and responded of external stakeholders on voluntary basis as well as among the consortium in line with the work plan of DoA.

4.2 How will this data be shared?

CERTH: Some of the data will be published, but this will be decided during the preparation and submission of the corresponding paper. Those data that will eventually be shared will meet the provisions and terms of the grant and consortium agreements.

UAveiro: Most of the data will be published, but this will be decided during the preparation and submission of the corresponding paper. Those data that will eventually be shared will meet the provisions and terms of the grant and consortium agreements.

ULEIC: Suitable data will be published, but this will be decided during the preparation and submission of the corresponding paper. The shared data will meet the provisions and terms of the grant and consortium agreements.

UNEW: Some of the data will be published, but this will be decided during the preparation and submission of the corresponding paper. The data that is eventually shared will meet the provisions and terms of the grant and consortium agreements.

NZ: Part of the data will be published/shared, but this will be decided during the preparation and submission of the corresponding paper. The data that will eventually be shared will meet the provisions and terms of the grant and consortium agreements.

CES: See 4.1.

MMU: Most of the data will be published, but this will be decided during the preparation and submission of the corresponding paper. Those data that will eventually be shared will meet the provisions and terms of the grant and consortium agreements.

YSD: The data that will be published in corresponding papers will meet the provisions and terms of the grant and consortium agreements.

ETA: Data will be shared through various channels, depending on its level of accessibility. For instance, publicly available data will be disseminated through project website, social media, and EU platforms (CORDIS, Open Research Europe), newsletter, conferences, and events. Meanwhile, Internal data will be circulated through file-sharing system among project partners.

TBWR: See 4.1.

4.3 Which are the main data access procedures?

CERTH: The data can be requested from their creators by interested project partners. For the data that will eventually be published, their access procedures will be based on the publication policy of the corresponding journal (e.g., gold open-access, green open-access, subscription model).

UAveiro: For the data that will be published, their access procedures will be based on the publication policy of the corresponding journal (e.g., gold open-access, green open-access, subscription model).

ULEIC: The access procedures for the data that will be published will be based on the publication policy of the corresponding journal (e.g., gold open-access, green open-access, subscription model).

UNEW: The data can be requested from their creators by interested project partners. For the data that is published, at least green open-access will be guaranteed and gold open access will be used when it is available.

NZ: For the data that will eventually be published, their access procedures will be based on the publication policy of the corresponding journal (e.g., gold open-access, green open-access, subscription model).

CES: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values

MMU: The data can be requested from their creators by interested project partners. For the data that will eventually be published, their access procedures will be based on the publication policy of the corresponding journal (e.g., gold open-access, green open-access, subscription model)

YSD: The data can be requested from their creators by interested project partners. For the data that will eventually be published, their access procedures will be based on the publication policy of the corresponding journal (e.g., gold open-access, green open-access, subscription model).

ETA: See 4.3 and 4.2.

TBWR: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

4.4 Which are the technical mechanisms for dissemination and the relevant software to access /reuse the data?

CERTH: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

UAveiro: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values

ULEIC: The dissemination will be made through journal/conference publications/ workshops/ talks. The data is independent of any software, ensuring universal accessibility and making interpretation intuitive.

UNEW: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

NZ: The dissemination of the agreed data will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

CES: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

MMU: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

YSD: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

ETA: Project website (external), LinkedIn platform (external), File-sharing platform (internal).

TBWR: The dissemination will be made through journal/conference publications/ workshops/ talks. The data are software independent as they are simple numerical values.

5. Data Archiving and Preservation (including storage and backup)

5.1 Which procedures will be put in place for data long-term preservation?

CERTH: Data will remain available in the IT infrastructure of the experiment partner(s) in the form of .txt files. Some of that data will be organized in the form of Excel files.

UAveiro: Data will remain available in the IT infrastructure of the experiment partner(s) in the form of .txt files. Some of that data will be organized in the form of Excel files.

ULEIC: Suitable data will be available through publications and stored in the Figshare repository. Data will remain available in the IT infrastructure of the University of Leicester. The files will be organized in .docx, .txt, .jpeg and .xls formats.

UNEW: Suitable data will be available through publications. Rest of the data will remain available in the IT infrastructure of the Newcastle University. Newcastle University provides Research Data Warehouse service that can be used to store data safely and securely. The files will be organized in .docx, .txt, .jpeg and .xls formats.

NZ: Data will remain available in the IT infrastructure of Novonesis in the form of .txt or .xls files.

CES: Publications and deliverables will remain available even beyond the project end.

MMU: Data will remain available in the IT infrastructure of the experiment partner(s) in the form of .txt files. Some of that data will be organized in .docx, .txt, .jpeg and .xls formats.

YSD: Data will remain available in the IT infrastructure of the experiment partner(s) in the form of .txt files. Some of that data will be organized in the form of Excel files.

ETA: Archival storage on project website: public reports, factsheets, publications and key dissemination materials will remain accessible on the project website for a specified period beyond the project's completion. Internal project repository: A secure repository (SharePoint) will store restricted-access and confidential materials.

TBWR: Publications and deliverables will remain available even beyond the project end.

5.2 How long should data be preserved and what is its approximated end-volume?

CERTH: The data will be preserved indefinitely. The approximated end-volume is not expected to exceed few Megabytes. Typical back-up practices will be applied.

UAveiro: The data will be preserved indefinitely. The approximated end-volume is not expected to exceed few Megabytes. Typical back-up practices will be applied.

ULEIC: The data will be preserved in the University IT system for at least 5 years. Peer-reviewed publications will be in open access Journals, which will be available indefinitely. The approximate end-volume is not expected to exceed a few Megabytes.

UNEW: The data will be preserved in the University IT system for at least 5 years. Peer-reviewed publications will be in open access Journals, which will be available indefinitely. The approximate end-volume is not expected to exceed a 50 Megabytes.

NZ: The data will be preserved indefinitely. The approximate end-volume is not expected to exceed a few Megabytes. Typical back-up practices will be applied.

CES: The approximated end-volume is not expected to exceed few Megabytes. Typical back-up practices will be applied.

MMU: The data will be preserved in the University IT system for at least 5 years. Peer-reviewed publications will be in open access Journals, which will be available indefinitely. The approximate end-volume is not expected to exceed a few Megabytes.

YSD: The data will be preserved indefinitely. The approximated end-volume is not expected to exceed few Megabytes. Typical back-up practices will be applied.

ETA: Public publications remain for an indefinite amount of time. Confidential data will be preserved as long as necessary and deleted when no longer relevant.

TBWR: The approximated end-volume is not expected to exceed few Megabytes. Typical back-up practices will be applied.

7. Open access provisions and sharing during the project

A dedicated budget is foreseen to ensure that publications stemming from the REUSE project fall under the “gold open access” regulation: free universal access will be given immediately upon publication for academics and the public alike to the final published versions of articles; all articles will be immediately provided in open access mode as published. Several scientific journals providing “gold open access” rights have been already identified by the Consortium and will be the preferred target for publication (e.g., IJGHG, Energy, Applied Energy etc.). In addition, each beneficiary will deposit a machine-readable electronic copy of the published

version or final peer reviewed manuscript accepted for publication in a repository for scientific publications. This step will be followed even in case of open access publishing (“gold open access”) in order to ensure long-term preservation of the article. The repository for scientific publications is an online archive and Open Access Infrastructure for Research in Europe (OpenAIRE) and will be the entry point for researchers to determine what repository to choose (<http://www.openaire.eu>). *For those publications which cannot follow the “gold open access” publication pathways (if any), a self-archiving “green open access” option will be adopted.*

Project data will be shared via a dedicated online repository in Microsoft TEAMS. Figure 1 illustrates the corresponding folders.

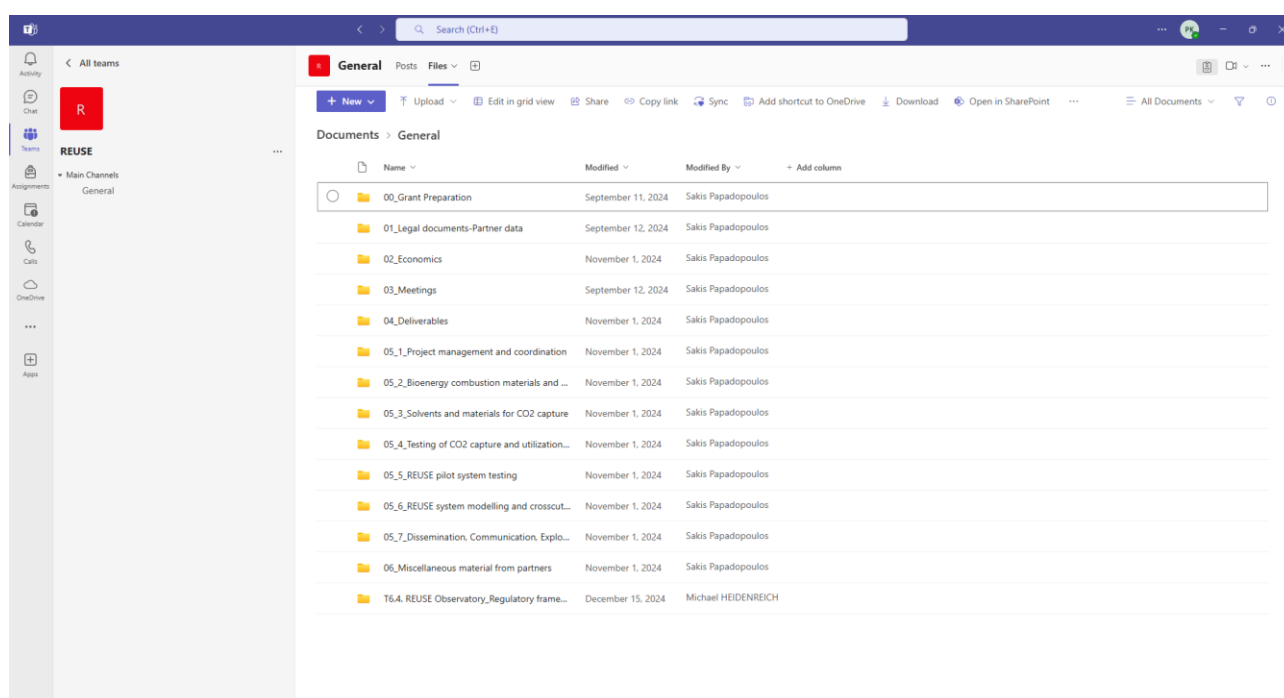


Figure 1: TEAMS repository for data sharing by partners

- 0 **Grant preparation:** This folder includes files that were used in the preparation of the grant agreement.
- 1 **Legal documents-partner data:** This folder includes data and information such as the grant agreement and the consortium agreement.
- 2 **Economics:** This folder includes economic data such as the pre-financing distribution of funds, the distribution of the intermediate and final funding etc.
- 3 **Meetings:** This folder includes all the necessary information for and from the project meetings, including agendas, travelling, technical and management presentations etc.
- 4 **Deliverables:** This folder includes the project deliverables, their draft and final versions.

- 5 **WPs:** This folder includes separate WP files that are shared among the partners to facilitate the progress of the work.
- 6 **Miscellaneous partner data:** This folder includes all the data that needs to be shared but does not fall under the above categories.
- 7 **REUSE Observatory:** D6.3 document is currently uploaded in this folder.

8. Provisions for data from surveys

The study of social acceptance of the developed technologies is included in Work Package 6 (WP 6). The REUSE Observatory will conduct questionnaires and surveys of professionals, researchers and public. Research activities in WP6 will investigate upon elements linked to the level of acceptance and perception of security measure, for different categories of end-users, paying particular attention to sex and gender-related aspects and other intersecting factors. Compliance with identified needs and requirements, usability, inclusiveness, accessibility, of the REUSE strategy will be investigated as well from a gender perspective. The collection of information related to sex and gender of research participants, in surveys (WP6) will be subject to a two-step method that “measures birth sex and current gender identity separately” will be implemented, as per the recommendations of the European Commission (EC, Gendered Innovations 2: How Inclusive Analysis Contributes to Research and Innovation Policy, July 2020, p. 192). REUSE research activities will allow for analyzing sex and gender, while also considering other intersecting factors, and examining similarities and differences between groups. Sex and gender distribution will be duly reported.

9. Ethics and Privacy

All project activities carried out in the REUSE project will be conducted in compliance with fundamental ethical principles. All the partners will process the data in a manner that would not hamper the guidelines set up by the EU in issues such as:

- Details on the procedures to be implemented for data collection, storage, protection, retention and destruction and confirmation that the project will comply with the relevant national and EU legislation.
- Details on the informed consent procedures to be implemented.
- Confirm that the existing data is publicly available, or otherwise provide relevant authorisation.

Privacy of the project participants will be secured. No person or organisation involved (e.g., in surveys) will be intentionally and unintentionally identifiable directly or indirectly in the datasets. To ensure compliance with data protection regulations and safeguard respondent privacy, all direct identifiers (such as names and contact details of individuals and organizations) will be stored separately from the data unless otherwise

contractually agreed. Additionally, any indirect references (e.g., specific business sectors, branches, or industries) will be removed after the anonymized dataset has undergone thorough validation and verification.

10. Partners

