



Funded by
the European Union



MEMCAT

Membrane-assisted Ethylene Synthesis over Nanostructured Tandem Catalysts

HORIZON EUROPE GRANT AGREEMENT NUMBER: 101130047

Start date of project: 01/05/2024

Duration: 4 years

D6.1 Data Management Plan

Call identifier	HORIZON-EIC-2023-PATHFINDEROPEN-01-01	
Deliverable:	D6.1 Data Management Plan	
Work Package:	WP6 Project Management	
Due date:	31/10/2024	
Submission date:	31/10/2024	
Organisation Responsible of Deliverable:	INL	
Version	1.1	
Author name(s):	Nuria Barros (INL)	
Reviewer(s):	Yury Kolenko (INL), Laura Salonen (UVIGO)	
Type:	<input type="checkbox"/>	R – Report
	<input type="checkbox"/>	Websites, patents filling, etc.
	<input checked="" type="checkbox"/>	DMP
	<input type="checkbox"/>	E – Ethics
Dissemination level:	<input checked="" type="checkbox"/>	PU - Public
	<input type="checkbox"/>	SEN - Sensitive limited under the conditions of the Grant Agreement

Version	DATE	Changes	CHECKED	APPROVED
V1	28-10-2024	First Version	Nuria Barros	Yury Kolenko
V2	31-10-2024	Final Version	Laura Salonen, Fausto Gallucci, Karoliina Honkala, Nuria Barros	Yury Kolenko

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

Content

1. Executive Summary	3
2. Introduction	3
3. Definitions.....	4
4. Data Summary	4
5. FAIR Data	8
5.1 Making data findable, including provisions for metadata	8
5.2 Making Data Openly Accessible	8
5.3 Making data interoperable	11
5.4 Increase data re-use (through clarifying licenses)	11
6. Allocation of Resources and Data Security.....	11
7. Conclusions and revision cycles	12

1. Executive Summary

The present document is the Deliverable 6.1 Data Management Plan within WP6 Project Management of the MemCat project.

The main objective of D6.1 is to provide the plan for managing the data generated, collected, and transferred during the project. Likewise, this document will include the necessary information on how the data will be preserved and handled as well as the methodologies used to accomplish the tasks. This first Data Management Plan (DMP) shows that there are mainly 12 data sources that will be produced as part of the project activities and that are relevant to be included in the DMP.

Following the EU's guidelines regarding the DMP, this is a living document and hence several issues will be updated during the project lifetime: in months 30 (31/10/2026) and 48 (30/04/28).

2. Introduction

Research and innovation projects such as MemCat usually produce large sets of data. Depending on the discipline, the data could come for example from laboratory testing, prototyping, or observations. However, what will happen with the data after it was collected, analyzed, and the project finished, often remains unclear and uncertain. Furthermore, a lot of data sets are potentially interesting for other researchers, but due to the fact that they are either stored on local servers or miss crucial metadata (or both), their potential value cannot be exploited. Hence, the purpose of the DMP is to identify and plan the data that will be produced at the beginning of the research project.

Herein we list datasets that will be generated and collected during MemCat and provide a plan for managing the data in an efficient way. The DMP also describes the standards that will be used, how the research data will be preserved, and what parts of the datasets will be shared for verification or reuse. By following the principle *as open as possible, as close as necessary*, we contribute to the reuse of research data, making it more open, collaborative, and closer to society.

The general guidelines are the following:

- The research data policy will be supervised by the coordinator INL
- The guidelines for open and confidential data will be defined in the DMP
- Non-sensitive research data will be available during and after the project for access and re-use in the Zenodo repository
- MemCat will use Microsoft Teams for information exchange
- New knowledge will be published after agreed IP protection in peer-reviewed scientific journals and made available through open-access mechanisms/research depositories of the academic partners

3. Definitions

Data Management Plan (DMP): a working document that outlines how all available datasets will be handled both during the active project phase and after its end.

Data Owners: the individuals or groups of individuals who are held accountable for a dataset and who have legal ownership rights to a dataset even though that dataset may have been collected, collated, or disseminated by another party.

Dataset: a collection of data created in the course of the MemCat project or secondary data that can be published with due permission of the user. Datasets are the fundamental unit of data management.

Metadata: Information about datasets stored in a repository. Metadata is the term used to describe the summary information or characteristics of a set of data. In general terms, this means the What, Who, Where, When, and How of the data.

Primary data: original data that has been uploaded by the user especially for the purpose in mind.

Secondary data: data that has been captured for another purpose than the purpose in mind. Secondary data that is being reused, usually in a different context.

4. Data Summary

The main reason to describe how the data will be handled during the project is to ensure that all relevant data is collected in a safe and systematic manner, and that relevant datasets are published and given open access to benefit society.

MemCat will produce large amounts computational calculations data as well as various data documenting the synthesis and characterization of catalyst materials and their catalytic testing, nanofiller synthesis and characterization and their incorporation into membranes and their testing, membrane reactor design and component modeling, reactor testing, etc.

Considering the multidisciplinary character of the project, there are several complementary sources of data that will be generated and collected during the project, comprising data from simulations, synthesis, and screening. To this end, 11 major data sources have been identified:

1. Data from the synthesis experiments of the catalyst materials
2. Data from the characterization of catalysts
3. Data from catalysis
4. Data from *operando* catalysis experiments
5. Data from modelling
6. Data from the synthesis of the nanofiller materials
7. Data from the characterization of nanofillers
8. Protocols for CMS membrane preparation
9. Data from the experiments of membrane separation
10. Data from the fabrication of the membrane reactor prototype
11. Data from the experiments on the membrane reactor prototype testing and validation

Experimental datasets: collection of data generated during the experimental activities related to the MemCat project (e.g., simulations of catalysts, synthesis procedures for the catalysts, characterization data of the catalysts, catalytic data collected with reference and synthesized catalyst materials, nanofiller and membrane preparation and testing, catalytic membrane reactor performance data). We have identified the datasets (Table 1) that will be produced, gathered, observed, and reused in the MemCat project.

Table 1. Datasets overview

No.	Data type	Origin	WP#	Format
1	Data from the synthesis experiments of the catalyst materials	Primary data, collected at INL	1	Excel (XLSX, CSV), Word (DOCX, RTF), PowerPoint (PPT)
2	Data from the characterization of catalysts	Primary data, collected at INL	1	Excel (XLSX, CSV), Word (DOC, RTF), PowerPoint (PPT), IMAGE (TIFF, JPEG), Origin (OPJ)
3	Data from catalysis	Primary data, collected at MPG	2	Excel (XLSX, CSV), Word (DOC, RTF), PowerPoint (PPT), IMAGE (TIFF, JPEG), Origin (OPJ)
4	Data from <i>operando</i> catalysis experiments	Primary data, collected at MPG	2	Excel (XLSX, CSV), Word (DOC, RTF), PowerPoint (PPT), IMAGE (TIFF, JPEG), Origin (OPJ)
5	Data from modelling	Primary data, collected at JYU	2	GPAW Input and Output files, Excel
6	Data from the synthesis of the nanofiller materials	Primary data, collected at UVIGO	3	Excel (XLSX, CSV), Word (DOCX, RTF), PowerPoint (PPT)
7	Data from the characterization of nanofillers	Primary data, collected at UVIGO	3	Excel (XLSX, CSV), Word (DOC, RTF), PowerPoint (PPT), IMAGE (TIFF, JPEG), Origin (OPJ)
8	Protocols for CMS membrane preparation	Primary data, collected at UNICAL	3	Excel (XLSX, CSV), Word (DOCX), PowerPoint (PPTX), IMAGE (TIFF, JPEG).
9	Data from the experiments of membrane separation	Primary data, collected at TU/e	3	Excel (XLSX, CSV), Word (DOCX), PowerPoint (PPTX), IMAGE (JPEG, PNG), Origin (OPJ)
10	Data from the fabrication of the membrane reactor prototype	Primary data, collected at TU/e	4	Excel (XLSX, CSV), Word (DOCX), PowerPoint (PPTX), IMAGE (JPEG, PNG), Origin (OPJ)
11	Data from the experiments on the membrane reactor prototype testing and validation	Primary data, collected at TU/e	4	Excel (XLSX, CSV), Word (DOCX), PowerPoint (PPTX), IMAGE (JPEG, PNG), Origin (OPJ)

Table 2 describes the dataset and the purpose of the data collection & data generation in relation to the objectives of the project. Additionally, it shows the data utility for clarifying to whom the data might be useful.

Table 2: Dataset description and utility

No.	Data type	Description & Purpose	Utility
1	Data from the synthesis experiments of the catalyst materials	<p>Description The data contain the conditions used in the synthesis and the work-up of the reactions and materials</p> <p>Purpose The data describes the procedure that was followed to obtain a specific catalyst material.</p>	The data, together with that of #2, will aid in the optimization of the syntheses to obtain high-quality catalysts, as well as scale-up efforts.
2	Data from the characterization of catalysts	<p>Description The data contain analytical data from the characterization of the prepared catalysts.</p> <p>Purpose The data will confirm the preparation of the desired material and its state after catalysis.</p>	The data will assist in choosing the best synthesis method for preparing high-quality material, to assist in the interpretation of the catalytic results based on the catalyst materials' properties before and after catalytic activity testing towards CO ₂ conversion to ET. The data will also assist in designing robust synthesis methods for the scale-up process.
3	Data from catalysis	<p>Description The data contains yields of the products obtained and the description of the measure conditions.</p> <p>Purpose The data will indicate the catalytic activity of the different materials.</p>	The data will be used to access the activity of the different prepared catalysts towards CO ₂ conversion to ET enabling the selection of the most active materials. Moreover, the data can be used to understand the kinetics of the reaction in the differently synthesized materials.
4	Data from <i>operando</i> catalysis experiments	<p>Description The data includes DRIFT, UV/Vis, and Raman spectroscopy results obtained.</p> <p>Purpose The data will show the physical properties of the catalyst and the reaction intermediates under different conditions.</p>	The data will be used to understand the active phase of the materials, the reaction intermediates and deactivation processes during the reaction.
5	Data from modelling	<p>Description The data contain the structural information of potential catalysts, the acid–base properties of the oxide support, thermodynamic and kinetic energies for CO₂ activation/hydrogenation, and computed spectra.</p>	The data will point out optimum compositions. At the same time, these data will allow the understanding and the search of descriptors.

		Purpose Anticipate and compare the relative ability of the catalysts.	
6	Data from the synthesis of the nanofiller materials	Description The data contain the conditions used in the synthesis and the work-up of the reactions and materials Purpose The data describes the procedure that was followed to obtain a specific nanofiller material.	The data, together with that of #7, will aid in the optimization of the syntheses to obtain high-quality nanofillers, as well as scale-up efforts.
7	Data from the characterization of nanofillers	Description The data contain analytical data from the characterization of the prepared nanofillers. Purpose The data will confirm the preparation of the desired material.	The data will assist in choosing the best synthesis method for preparing high-quality nanofiller material and in the interpretation of the membrane separation results based on the nanofiller materials' properties. The data will also assist in designing robust synthesis methods for the scale-up process.
8	Protocols for CMS membrane preparation	Description The data contain the conditions used in the preparation of the CMS membranes. Purpose The data describes the procedure that was followed to obtain a specific membrane.	The data, together with that of #9, will aid in the optimization of the membrane preparation to obtain high efficiency in ET separation.
9	Data from the experiments of membrane separation	Description The data contain analytical data from the gas separation tests on the membranes. Purpose The data will confirm the efficiency in gas separation of the desired material.	The data will assist in choosing the best nanofiller material and the preparation method for an efficient ET separation membrane.
10	Data from the fabrication of the membrane reactor prototype	Description The data contain the procedures used in the development of the membrane reactor prototype. Purpose The data describes the procedure that was followed to obtain the membrane reactor prototype.	The data will allow, together with that of #11, the fabrication of a high-performance membrane reactor prototype for efficient ET production.
11	Data from the experiments on the membrane reactor prototype testing and validation	Description The data contain analytical data from the membrane reactor prototype testing and validation. Purpose The data will confirm the efficiency of the membrane reactor prototype to produce ET.	The data will allow the analysis and evaluation of the long-term performance of the membrane reactor prototype.

5. FAIR Data

5.1 Making data findable, including provisions for metadata

General data will be standardized in a way that it can be understood by the partners, allowing for successful transfer of knowledge. To this end, a metadata file will be created that will summarize the major characteristics of the dataset and linked within each dataset when data is collected and shared. The MemCat consortium will generate and update version of the guidelines and a document template for establishing consistent metadata among the different datasets generated. This metadata record will allow correct identification and suitable reuse/reprocessing of the dataset.

The data will be placed in a folder containing a metadata file indicating at least:

- The title of the project and the name of the data (identifier dataset)
- Author, date, and partner that has generated the data
- The WP (and the task) related to that data
- Description of the methodology and analysis used to generate the data
- If newer versions are uploaded, this will be indicated in the metadata file

Search keywords will be provided to optimize the possibilities for data re-use. Versions will be clearly indicated. The details of the metadata will be agreed on in the next General Assembly.

The MemCat publications will receive DOI according to publishers' procedures and can be found by regular search engines. We will endeavor to publish datasets alongside these articles in the supplementary information and/or Zenodo repository. For modelling data, we will use [IDA fair data service](#).

We will create a Zenodo community for MemCat and ICUBE will manage the data to be uploaded.

5.2 Making Data Openly Accessible

The following table provides data that are produced and used in the project and that will be made openly available. It also explains why several datasets cannot be shared. For these cases, an alternative solution is provided.

Table 3: Dataset accessibility

No.	Data type	Data openly available	Justification	Alternative solution
1	Data from the synthesis experiments of the catalyst materials	Yes	After optimum synthesis conditions have been found and the possibility of IP protection evaluated, the successful synthesis conditions will be published.	
2	Data from the characterization of catalysts	Yes	After optimum synthesis conditions have been found for high-performance catalysts and the possibility of IP protection evaluated, the characterization data will be published.	
3	Data from catalysis	Yes	After defining the catalytic activity and the possibility of IP protection evaluated, the successful catalytic systems will be published.	

4	Data from <i>operando</i> catalysis experiments	Yes	After identifying the reaction mechanism, the most important active sites, and the possibility of IP protection has been evaluated, the successful catalytic systems will be published.	
5	Data from modelling	Yes	After comparing with experimental results, the computational data will be published.	
6	Data from the synthesis of the nanofiller materials	Yes	After optimum synthesis conditions have been found and the possibility of IP protection evaluated, the successful synthesis conditions will be published.	
7	Data from the characterization of nanofillers	Yes	After optimum synthesis conditions have been found for high-performance nanofillers and the possibility of IP protection evaluated, the characterization data will be published.	
8	Protocols for CMS membrane preparation	Yes	After optimum conditions have been found for the membrane preparation and the possibility of IP protection evaluated, the preparation protocols will be published.	
9	Data from the experiments of membrane separation	Yes	After optimum conditions for the membrane separation have been found and the possibility of IP protection evaluated, the membrane separation data will be published.	
10	Data from the fabrication of the membrane reactor prototype	No		
11	Data from the experiments on the membrane reactor prototype testing and validation	Yes	After several test runs with different catalyst materials and membranes with high performance and the possibility of IP protection evaluated, the data obtained will be published.	

As indicated above, the following datasets will be made openly accessible: Data no #1–9 and #11. The following table describes the accessibility details of these particular datasets.

Table 4. Details on accessible datasets

No.	Data type	Location	Level of accessibility	Type of availability and required software tools	Information on metadata and additional data information
1	Data from the synthesis experiments of the catalyst materials	Zenodo repository	Public	Selected primary data; can be accessed with Word, application that displays pdf files	Metadata will be deposited in Zenodo
2	Data from the characterization of catalysts	Zenodo repository	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo
3	Data from catalysis	Zenodo repository	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo and Data archive at the Fritz Haber Institute
4	Data from <i>operando</i> catalysis experiments	Zenodo repository	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo and Data archive at the Fritz Haber Institute
5	Data from modelling	IDA Fair data service	Public	Selected primary files	Metadata will be deposited in Zenodo; data on IDA
6	Data from the synthesis of the nanofiller materials	Zenodo repository	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo
7	Data from the characterization of nanofillers	Zenodo repository	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo
8	Protocols for CMS membrane preparation	Zenodo repository	Public	Selected primary data; can be accessed with Word, application that displays pdf files	Metadata will be deposited in Zenodo
9	Data from the experiments of membrane separation	Zenodo and 4TU.ResearchData repositories	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo and 4TU.ResearchData

11	Data from the experiments on the membrane reactor prototype testing and validation	Zenodo and 4TU.ResearchData repositories	Public	Selected primary data; can be accessed with Excel, Origin, application that displays pdf files	Metadata will be deposited in Zenodo and 4TU.ResearchData
----	--	--	--------	--	---

The datasets, papers, reports, and presentations, which have been decided to give open access to, will be published in the Zenodo repository. The consortium will follow the conditions, rules, and regulations from the Zenodo repository – including the settings for accessing the dataset. The software needed to access the data comprises Word, Excel, Adobe Acrobat, Origin, etc. MemCat will not have a formal data access committee.

5.3 Making data interoperable

The data produced in MemCat will be interoperable, allowing data exchange and re-use between researchers, institutions, organizations, countries, etc. MemCat will, whenever possible, adhere to standards for formats and as much as possible comply with available (open) software applications.

5.4 Increase data re-use (through clarifying licenses)

MemCat will publish the data as soon as IP protection has been clarified, IP has been protected when necessary, and the data has got clearance for publication. Open access to all peer-reviewed publications will be assured. The relevant data will be made available after 6 M from the acceptance of the publication the latest. Further aspects of data re-use by Third parties will be discussed in the updated DMP (M30).

6. Allocation of Resources and Data Security

The data produced within MemCat will be securely stored and regularly backed up through a **shared Repository**: MemCat project uses Microsoft Teams for sharing documents, administrative matters, minutes, reports, and deliverables. INL is responsible for the overall maintenance of the document repository as well as for the uploading of documents, but every WP leader is obligated to take charge of their respective WP folders.

The consortium will use the free-of-charge Zenodo repository for making datasets accessible, and the IDA repository for dataset #5, ensuring that the data are safely stored in this certified repository for long-term preservation and curation. The handling of the Zenodo repository on behalf of MemCat as well as all data management issues related to the project fall in the responsibility of the coordinator. Each partner is responsible for adding the datasets.

As for the publications, the consortium will publish them in scientific journals that allow open access. The costs related to open access will be claimed as part of the grant.

7. Conclusions and revision cycles

The Data Management Plan is currently under implementation and includes all datasets identified in Section 4. The MemCat consortium is responsible for verifying the application of rules stated in the present document. This document will have the next review by month 30 (31/10/2026). The upcoming DMPs will provide updates on the issues raised above and more information on the following aspects: metadata, data interoperability; data re-use by Third parties.