SOMMER

Solar-Based Membrane Reactor for Syngas Production

D1.4 Data Management Plan

WP1 – Project Management, Coordination and Dissemination

30.04.2024

Disclaimer

The work described in this document has been carried out within the framework of the SOMMER project, generously funded by the European Union. The views and opinions expressed herein are the sole responsibility of the author(s) and do not necessarily represent those of the European Union or the granting authority, the European Climate, Infrastructure and Environment Executive Agency (CINEA). The European Union and CINEA cannot be held accountable for any interpretations or applications of the information contained in this document.

Acknowledgement

The project SOMMER is funded by the European Union under Grant Agreement Number 101118293.

Document Identifier

Grant Agreement No.	101118293		Acronym	SOMMER
Title	Solar-based membrane reactor for syngas production			
Call Topic	HORIZON-CL5-2022-D3-02-06			
Type of action	HORIZON EUROPE, RIA - Research and Innovation Actions			
Start date	01.11.2023		Duration	48 months
Project URL	https://www.project-sommer.eu			
Project Officer	Dr. Luca Bondi			
Project Coordinator	DLR - German Aerospace Center			
Deliverable	D1.4 Data Management Plan			
Date of Delivery	Contractual	30.04.2024	Actual	30.04.2024
Nature	Report	Dissemination level		Public
Lead Beneficiary	DLR - German Aerospace Center			
Responsible Author	Asmaa Eltayeb (DLR), asmaa.eltayeb@dlr.de			
Contributors	Nicole Neumann (DLR), Nicole.Neumann@dlr.de			
Reviewers	Maria Balaguer Ramirez (CSIC), mabara@itq.upv.es Stefan Baumann (FZJ), s.baumann@fz-juelich.de			
Final version approval date	30.04.2024			
HISTORY OF CHANGES				
VERSION	PUBLICATION DATE	CHANGE		
1.0	14.03.2024	Initial version		

About the Project

SOMMER aims to develop and demonstrate an innovative carbon-neutral process for syngas production by directly integrating solar energy into a catalytic membrane reactor, facilitating the decomposition of H_2O and CO_2 (e.g., captured from carbon-emitting industries or through direct air capture). This approach enables SOMMER to overcome reliance on fossil-based energy for syngas production, utilizing CO_2 instead of natural gas as a feedstock. Syngas, a critical intermediate for the chemical industry, prompts SOMMER to encompass the entire value chain - from CO_2 provision in a cement plant to syngas formation and further processing into valuable products like DME or methanol. The core of SOMMER's technology is the optimized energy integration of a novel thermochemical conversion process of CO_2 and H_2O in a single step. This process is supported by highly selective catalysts, an oxygen transport membrane, and a concentrated solar-thermal plant fulfilling the thermal energy demand. The key outcomes of SOMMER involve the experimental demonstration and evaluation of the innovative solar-powered membrane technology. Additionally, it focuses on developing high-performance, cost-effective membranes as pivotal components, elevating the technology to new heights. SOMMER's strategy involves advancing membrane manufacturing through slip-casting or extrusion, more mature approaches, and additive manufacturing to optimize the effective membrane surface area in the reactor. The concept anticipates future advantages, allowing prolonged and flexible operation by seamlessly switching between two operational cases: I) Purely solar approach at 1500 °C and II) a biogas-supported approach at 900 °C. Furthermore, SOMMER aims to identify the technological, ecological, and economical potential for flexible and highly efficient solar syngas production, contributing to the formulation of a detailed roadmap and providing a foundation for pre-commercialization through subsequent R&D development activities.

DLR	Deutsches Zentrum Für Luft - und Raumfahrt e.V.	DE	
FZJ	Forschungszentrum Jülich GmbH	DE	
IREC	Fundacio Institut De Recerca De L'Energia De Catalunya	ES	
HTE	HTE GmbH The High Throughput Experimentation Company	DE	
CSIC	Agencia Estatal Consejo Superior De Investigaciones Científicas	ES	
MAM HW	Morgan Advanced Materials Haldenwanger GmbH	DE	
TITAN	TITAN Cement Company S.A.	GR	
BASF*	BASF SE	DE	

*Associated Partner

Document Summary

The document represents the inaugural edition of Deliverable D1.4, the 'Data Management Plan', developed within WP1 of the HORIZON Europe SOMMER project. It outlines the creation of the Data Management Plan (DMP) for the data to be generated and utilized within the project's scope. This initial version of the DMP, formulated early in the project timeline, serves as a foundational framework to delineate the data management processes to be adhered to by each partner, as well as to outline the overarching data management strategy at the project level.

This DMP will undergo two revisions, conducted as part of the interim and final project reviews, to align with the evolving data landscape and the identified uses by the consortium. As part of the effort to ensure research data compliance with FAIR principles (i.e., making data findable, accessible, interoperable, and reusable), D8.2 includes comprehensive information and clear descriptions pertaining to:

a) The types of datasets to be generated, collected, and processed;
b) The primary criteria utilized to categorize data as (1) private, (2) publishable, or (3) Open Access;
c) The methodologies and standards adopted for data processing;
d) The consortium's policy concerning data sharing;
e) The procedures for data handling, archiving, and preservation throughout and beyond the project, alongside the metadata generated to manage the lifecycle of collected datasets.

Changes with Respect to the DoA

No changes or deviations from the work plan have occurred with respect to this Deliverable D1.4, the Data Management Plan.

Table of Contents

1. Introduction

The Deliverable D1.4,'Data Management Plan', is developed within WP1 Project Management, Coordination, and Dissemination of the SOMMER project. Its purpose is to outline the approach to be adopted during and after the project for relevant data acquisition, organization, and sharing. The document is based on the Description of Work (DoW) in the proposal and its modifications during the Grant Agreement preparation procedure. It is drafted using the template provided by the Horizon Europe online manual.

In this context, the specific objectives of this deliverable are as follows:

- Identify the datasets that the project will generate as it progresses and categorize them based on format and content.
- Establish processes to ensure that the data adheres to the principles of being 'FAIR' (Findable, Accessible, Interoperable, and Reusable).
- Ensure robust data management practices in compliance with FAIR principles and the General Data Protection Regulation (GDPR).
- Establish procedures for the handling, archiving, and retention of data during and after the project.
- Draft and refine the consortium's data sharing policy.

The DMP will continue to evolve over the life of the project reflecting new input into data management. According to the Grant Agreement, the DMP is intended to be revised twice, follow-up versions shall follow as part of the interim and final project review.

2. Data Summary

The **types, formats, and sources of data** that the project will generate/collect encompass both experimental and numerical research data. These data will be generated by the partners themselves during the project, originating from:

- Laboratory tests and characterizations,
- Synthesis of oxide multi-component powders and their shaping into membrane tubes with various morphology and microstructure,
- Modeling, operation simulation, and relevant proof-of-concept demonstration campaigns of reaction management of redox oxide catalytic membrane and green chemical production,
- Techno-economic analysis (TEA) and life cycle assessment (LCA) tasks.

This data will include, among other things, literature-based identification of suitable material for the membranes and catalysts, synthesis/shaping protocols, thermophysical/thermomechanical properties, kinetics/thermodynamics data, process date from solar experiments, process simulation and automation outputs, and TEA & LCA data.

The specific origin of the data will vary depending on the dataset. Table 1 provides an initial indicative list of the types of datasets expected to be generated and managed throughout the evolution of the SOMMER project, along with their origin and the partners expected to generate them.

Table 1: First classification and indicative examples of types of datasets to be created and managed within the SOMMER project from project partners

Dataset/research output	Exemplary software	Data type
Text (e.g. reports, documents)	Word, Latex	*.txt, *.docx, *.doc, *.tex, *.pdf (PDF/A-1), *.html
Presentations	PowerPoint, Latex	*.pptx, *.pdf
Graphics/Diagrams	Excel, Visio, Inkscape, PowerPoint, Python, etc.	*.svg, *.png, *.tif, *.jpeg, *.pdf, etc.
Spreadsheets (calculation tables, experimental data)	Excel, Python	*.csv, *.xls, *.xlsx, *.py, *.ipynb
Audio	vlc	*.wave, *.mp3
Video	vlc	*.avi, *.mp4
CAD files of reactor/test rigs designed	Autodesk Inventor	*.ipt, *.iam, *.idw, *.stp, *.iges, *.stl, *x_t, *x_b etc.
CFD & FEM Analysis Files	OpenFOAM, Ansys, COMSOL	*.csv, *.json, ascii
LCA data files	OpenLCA, Python	*.JSON, *.py, *ipynb
CFD & FEM Analysis Files	OpenFOAM, Ansys, COMSOL	*.csv, *.json, ascii

Purpose of data collection/generation and its relation to the objectives of the project: All these groups of data to be collected are necessary and will be used for the successful implementation of the various research tasks of the project. Specifically:

- Material characterization data will refine understanding of synthesized oxides' properties and their variations under different conditions, also serving as a benchmark for theoretical predictions.
- Synthesis recipes and shaping protocols will establish reliable, scalable, and eco-friendly methods for material synthesis.
- CAD, graphics, P&ID, and automation data will contribute to the design and operation of the project's central deliverable, the solar membrane reactor unit.
- Process simulation data will validate experimental results and refine simulation models.
- LCA data will guide eco-friendly materials selection and the ecological assessment of the overall process chain. TEA will assess overall performance against existing electrolysis and thermochemical systems.

Furthermore, these datasets will support transversal project activities such as communication, dissemination, exploitation strategy formulation, intellectual property rights (PR) management inside the consortium, and overall project management.

Expected data size: The project anticipates generating < 1 TB of data, with the figure subject to refinement as the project progresses.

Reuse of existing data: TEA and LCA activities may leverage previous methodologies or templates related to the project's research areas to streamline relevant work.

Data utility: Clearly, the data generated will directly benefit consortium partners, guiding their efforts towards achieving the project's technical objectives. Additionally, sensitive data will be utilized to identify Key Exploitable Results and develop pertinent Intellectual Property Rights (IPR) strategies.

Moreover, data published in scientific journals and presented at conferences will prove valuable to the broader research and industrial communities engaged in relevant fields such as materials science and engineering, oxide ceramics synthesis and shaping, engineering process design and implementation, solar/renewable energy, etc. Overall, the generated data will be useful to the first list of main potential stakeholders and industrial end users identified in the "Impact" section of the Project, namely:

- Shaped ceramics industry
- Petro-chemical
- Industrial processes with high CO_2 emission
- CSP plants developers
- Political governmental authorities
- Public Communities.

3. FAIR data

3.1. Making data findable, including provisions for metadata

Measures and procedures to ensure the discoverability of SOMMER project data include providing accompanying datasets with properly structured and accurate metadata in accordance with FAIR data principles, and utilizing standard identification mechanisms where applicable. For the latter, persistent identifiers such as Digital Object Identifiers (DOIs) or Uniform Resource Names (URNs), along with ORCID researcher identifiers, will be employed.

Various datasets will be linked to corresponding DOIs. For instance, DOIs for scientific publications will be provided by the publisher, while DOIs for other literature (e.g., reports) will be assigned by the repository where they will be archived. Published and FAIR-compliant data will be archived in trusted open data repositories, such as Zenodo, a collaborative effort of CERN, the OpenAIRE initiative, and the European Commission (https://zenodo.org/). Within Zenodo, all uploads and published datasets are accompanied by standard metadata, including:

- Digital Object Identifiers
- Version numbers
- Bibliographic information
- Keywords
- Abstract/description
- Associated project and community
- Associated publications and reports, specifying types (e.g., Publication, Poster, Presentation, Dataset, Video/Audio) with mandatory basic information: publication date, title, authors, and abstract
- Funding details: Grant programme name and number
- Access and licensing information: access rights (open, embargoed, restricted, closed access), license name (Creative Commons versions)
- Language

Additionally, links to all project uploads will be provided to the Open Access Research Data repository initiated by the European Commission (OpenAIRE https://www.openaire.eu/). Each partner will be responsible for uploading public datasets they have generated and assigning specific relevant keywords to these datasets.

3.2. Making data openly accessible

The project partners will make every effort to maximize the accessibility and usability of project-generated data for third parties, particularly for study, teaching, and research purposes. All reports and data related to the project's Public (PU) Deliverables, which are not deemed confidential, will be made available as a general principle. However, exceptions and potential embargo periods, particularly relating to scientific publications, protection of partners' IPR, and potential commercial exploitation, will be determined by the consortium as outlined in the respective Grant Agreement and Consortium Agreement. Additionally, certain exceptions apply to findings from eleven non-public, Sensitive (SEN) Deliverables, identified as pertinent for potential exploitation by the partners, as explicitly stated in Table 2 below.

Table 2: Sensitive, non-public Deliverables and respective reason

Deliverable	Deliverable name	Reasons for exception
D1.5	Dissemination and exploitation plan	Contains listing of stakeholders, other sensitive data
D3.1	Definition of operational parameters of solar membrane reactor operation (pressure and temperature range), as well as gas compositions.	Exploitation potential, patent application
D3.2	Selection of most advantageous approach for pO2 reduction for Case I operation	Exploitation potential, patent application
D2.1	Identification of suitable material system for membranes and catalysts	Exploitation potential, patent application
D2.2	Membrane manufacturing per slip-casting in lab-scale size using the most promising identified material candidates established	Strong exploitation potential, commercial application
D2.3	Membrane manufacturing per 3D-printing in lab-scale size using the most promising identified material candidates established	Strong exploitation potential, commercial application
D2.4	Oxygen permeability data of (first generation) membrane samples ready	Exploitation potential, partly publication in form of journal articles planned
D2.5	Definition of multi-layer catalytic membrane material and microstructure for scaled-up membrane units for operation in the solar membrane reactor	Exploitation potential, partly publication in form of journal articles planned
D4.1	Model of solar membrane module completed and verified	Exploitation potential, partly publication in form of journal articles planned

D2.6	Kinetics measurement data of most promising catalytic membrane assemblies collected	Exploitation potential, partly publication in form of journal articles considered, commercial application, patent application
D4.2	Design of solar interface completed	Core technology output of the project, possible patent application, commercial potential, interest from the industrial partner for application in H2 production
D1.7	Dissemination and exploitation plan updates	Contains listing of stakeholders, other sensitive data
D5.1	Model of solar membrane reactor design finalized	Core technology output of the project, possible patent application, commercial potential, interest from the industrial partner for application in H2 production
D6.1	Design parameters for CO_2-providing process (cement plant) defined for process simulation	Contains protected background information of the industrial partner
D5.2	Solar flux guide tested	Exploitation potential, partly publication in form of journal articles planned
D3.4	Calculate cost range of end product methanol / DME	Contains protected background information of the industrial partner, partly publication in form of journal articles planned

Furthermore, additional datasets beyond those listed in Table 2 may be identified during their generation as potentially holding commercial value for the partner(s) responsible for their creation, warranting additional protection measures. Examples of such datasets, though not exhaustive, include CAD and structural analysis files utilized in the design of solar membrane reactors, as well as process simulation files associated with them.

To ensure the easy discoverability and accessibility of project results and their associated metadata, it is planned to archive the data in the Zenodo open data trusted repository. Zenodo is a widely used platform for sharing research outcomes from projects funded by the European Commission, adhering to the FAIR principles. Within Zenodo, publications and other research outputs can be categorized under open, closed, or embargoed access, allowing for delayed sharing of the latter after a specified period.

In addition to Zenodo, some partners may opt to utilize their own research publication open repositories. For instance, DLR advocates for the Open Access principle through its own open current research information system ELIB (Electronic Library System), where mandatory self-archiving of all personnel publications is enforced. Publications are made publicly available and accessible through external searches, in accordance with the Green Open Access model, sometimes subject to a brief embargo period based on publisher agreements. ELIB can be accessed at https://elib.dlr.de/.

3.3. Making data interoperable

The interoperability of the generated data will be guaranteed by adhering to commonly endorsed metadata vocabularies, standards, formats, and methodologies, as well as by depositing the data in a repository that aligns with these vocabularies. As previously mentioned, the Zenodo repository is the likely solution, as it allows for the export of its internal metadata representation into other popular formats. In instances where it is necessary to use unavoidable ontologies or generate project-specific ones, these will be openly published alongside the open datasets to facilitate transparency and accessibility.

3.4. Increase data re-use

All reports and non-confidential data will be licensed under Creative Commons licenses CC-BY-SA or CC-BY, commonly used for open publications and relevant datasets, to ensure secure re-use of data and research outputs. These licenses will also include relevant disclaimers of liability for such re-use. Project partners will furnish the necessary documentation required to validate data analysis and facilitate data re-use, primarily through publications and as part of publicly available datasets on the Zenodo repository. Upon their availability as Open Access in these sources, data can be utilized by third parties, even after the conclusion of the project. As previously mentioned, in cases where scientific publications are expected to be derived from specific data, an embargo period will be enforced until acceptance by the publisher, including relevant documentation for data analysis and re-use. Similar restrictions regarding immediate public availability will be imposed on data identified as having potential commercial value, as mentioned in Section 3.2.

4. Allocation of resources

During the project, consortium partners will assume responsibility for securely managing datasets within their possession. Each partner will appoint *a member of its research team* involved in the project *as responsible for data management* and data assurance for their organization. Similarly, a *Data Manager* will be appointed to oversee data archiving and preservation during and beyond the project, managing the lifecycle of the generated datasets. Details of **the appointed individuals will be provided in the next revised version of the Data Management Plan (due M21).**

The Gold Open Access model is widely anticipated to be utilized by project partners when submitting manuscripts to journals. To accommodate relevant fees (ranging from 2,000 to 3,000 euros per published paper) from various publishing houses, budgets have been allocated for all research institutes and university partners within the consortium, with an allowance of 2 to 3 publications per partner. Notably, two German partners (including the coordinator) benefit from a nationwide agreement (DEAL) covering fees for Open Access publications with major academic publishers such as Wiley and Springer. A similar program is available for CSIC (PROA agreement). Additionally, exploring publication avenues in Open Access Journals without associated costs is also under consideration.

Industrial partners within the consortium may utilize a portion of their project budget to file any necessary patents to protect their intellectual property. Even when fees for this model are not budgeted in projects, DLR's (the coordinator) library can assume them under certain circumstances. The Creative Commons license CC-BY4.0 is recommended. DLR endeavors to retain copyright to articles and publications to enhance their reuse. DLR supports the Open Access principle through its own open

current research information system ELIB (Electronic Library System), where mandatory self-archiving of all personnel publications is enforced, sometimes after a short embargo period, in accordance with the Green Open Access model. These publications are publicly available and accessible through external searches. DLR considers collaboration as a core element of the research process and promotes it within the organization and among consortium partners in this and other international projects.

The Zenodo open data trusted repository intended for archiving the generated data is free of charge. Additionally, the TeamSite created by the Project Coordinator, DLR, within its own intranet domain serves as a communication and information exchange platform among partners, functioning as a data repository as well. All partners have been granted password-protected access and the capability to upload/download documents that can be stored and exchanged therein. As part of the coordinator's internal project management system, this "TeamSite-repository" does not incur any extra costs for the project or its partners.

The resources and terms for the long-term preservation of data will be deliberated among consortium partners at a later stage of the project. These discussions will encompass maintenance of the project's website (including public deliverables) and social media accounts, as well as the sharing of data deposited after the official end date of the project. Outcomes of these discussions will be included in subsequent, updated versions of the Data Management Plan.

5. Data security

During the project, consortium partners will bear responsibility for securely managing datasets within their possession. Each partner operates its own security system, typically relying on local, protected servers supplemented by additional backup services. Data hosted on the project's public website, www.project-sommer.eu, managed by DLR, and those on the TeamSite also managed by DLR, are subject to the respective hosts' data security, backup, and recovery mechanisms. With respect to long-term preservation and curation spanning after the project's end, as already mentioned it is currently envisaged that data will be archived in the Zenodo open data trusted repository.

6. Ethics

There are presently no ethical considerations identified within the SOMMER project. The technology under development in this project does not possess dual-use potential, thus export control will not impact the ability to publish data with Open Access. Should activities involving stakeholders entail questionnaires or other forms of contact and information exchange with the intention of making pertinent data publicly available, appropriate anonymization techniques will be implemented before data archiving. All dissemination activities related to the project, including the distribution of project material or information to third parties, will be planned to ensure a priori compliance with the provisions and requirements of the EU General Data Protection Regulation (GDPR).

7. Conclusion

The present Deliverable comprises the first version of the Data Management Plan for the SOMMER project, envisaged to evolve during the lifetime of the project in order to include renewed insights in the data management.

At first, the datasets that the project will generate through its course are identified and grouped following the evolution of the project. Then, the processes to be followed for rendering these data FAIR' (Findable, Accessible, Interoperable and Reusable) are established and procedures for data handling/archiving/preservation during and beyond the project are discussed.

All research and academic partners are strongly committed to promote open science policies, principles and research practices, encouraging their researchers to publish results and interpretations of their research in an open and transparent manner, yet respecting on the other hand confidentiality when this is required due to legitimate obligations (for example, with respect to the findings of 16 non-public deliverables identified as relevant for potential exploitation from the partners). In parallel, the consortium has appropriate technical and organizational skills and measures in place for data protection during the project. All partners are also committed to provide relevant metadata and their documentation to facilitate data's re-use and interoperability.