

HORIZON EUROPE PROGRAMME
TOPIC HORIZON-CL5-2021-D3-03-03

GA No. 101083700

**HYBRID TANDEM CATALYTIC CONVERSION PROCESS
TOWARDS HIGHER OXYGENATE EFUELS**



E-TANDEM - Deliverable report

Deliverable 1.2 – Initial Data Management Plan



**Funded by
the European Union**



Deliverable No.	E-TANDEM D1.2	
Related WP	1	
Deliverable Title	Initial Data Management Plan	
Deliverable Date	2023-06-30	
Deliverable Type	DMP	
Dissemination level	Public (PU)	
Author(s)	Gonzalo Prieto (CSIC) Irene Lamme (UNR)	2023-06-15
Checked by	Alessandra Lucini Paioni (UNR)	2023-06-15
Reviewed by (if applicable)	All partners	2023-06-21
Approved by	Gonzalo Prieto (CSIC) – Project coordinator	2023-06-28
Status	Final	2023-06-28

Document History

Version	Date	Editing done by	Remarks
V01	01/06/2023	UNR/CSIC	
V1.0	05/06/2023	UNR	
V1.1	08/06/2023	UNR	
V1.2	20/06/2023	CSIC	
V2.0	21/06/2023	All	
V3.0	28/06/2023	UNR	



Summary

Deliverable report D1.2 “Initial Data Management Plan” is a deliverable associated with the management work package via task 1.3 ‘Open Science Practices and data management’. In this report, the initial Data Management Plan (DMP) of E-TANDEM is presented, specifying the type of data that will be generated in the project, the standards and formats that will be used, how the data will be exploited and/or made accessible for verification and re-use, for the IPR protection strategy, and how this data will be curated, stored and preserved. This DMP is a living document, which will be updated throughout the lifetime of the project. A final DMP will be delivered at the end of the project (M42).

Within the E-TANDEM project, a large amount of data will be generated connected to production of carbon-neutral higher oxygenate e-fuels (HOEF). All work packages are involved in these activities. These data include output files from several physicochemical material characterization methods, microscopy techniques, process virtual layouts and simulation results, as well as experimental data from the demonstration of the (electro)catalytic conversions at laboratory/bench scale. Therefore, a DMP is necessary and vital to outline how the different project data will be handled during and after the E-TANDEM project focusing on the data collection methodology and data sharing issues.

This plan outlines the data that will be collected and generated during the project, how it will be stored and managed and the measures taken to ensure it will be accessible, assessable, usable, and interoperable by the project partners. It also describes the relevant data security. This DMP is based on the template provided by the European Commission [1] and will be updated and re-issued at intervals throughout the duration of the project as more information on the details of the data and its management become available to make sure the DMP remains current and relevant.



Contents

1	Introduction.....	6
1.1	E-TANDEM summary	6
1.1.1	Document structure	6
2	FAIR Data and E-TANDEM	7
2.1	Making data findable.....	7
2.2	Naming the data and files	7
2.3	Making data openly accessible.....	8
2.4	Making data interoperable.....	8
2.5	Increase data re-use (through clarifying licences)	9
3	Data Summary	10
3.1	Types of data and formats.....	10
3.2	Data Utility.....	11
4	Resource allocation for FAIR data and data security	12
4.1	Allocation of resources.....	12
4.2	IPR Protection Strategy	12
4.3	Data security.....	12
4.4	Other ethics and issues: use of external procedures for data management.....	13
5	Conclusion and recommendations.....	14
6	Risks and interconnections.....	15
6.1	Interconnections with other deliverables	15
7	Deviations from Annex 1	16
8	References.....	17
9	Acknowledgement.....	18



List of Tables

Table 2.1	Standard table format	11
-----------	-----------------------------	----

Abbreviations & Definitions

Abbreviation	Explanation
ASCII	American Standard Code for Information Interchange
CSV	Comma-Separated Values
DMP	Data management plan
DOI	Digital Object Identifier
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable and Reusable
FID	File IDentifier
GA	General Assembly (meetings)
HTML	HyperText Markup Language
IP	Intellectual Property
IPR	Intellectual Property Rights
JPEG	Joint Photographic Experts Group
OpenAire	Open Access Infrastructure for Research in Europe
PNG	Portable Network Graphics
PDF	Portable Document Format
RTF	Rich Text Format
TIFF	Tagged Image File Format
TXRM	Tomographic X-ray Microscopy
XRM	X-ray Microscopy

1 Introduction

The European Commission (EC) urges on the implementation of open access to scientific publications and research data derived from the European Union (EU) funded research projects.[1] This is reflected in the standardization of procedures to achieve Findable, Accessible, Interoperable and Reusable (FAIR) research data. The main objective behind this EC action is to maximise the impact of EU funded research projects and accelerate findings through the interaction of the data produced in different research projects. This document constitutes the initial Data Management Plan (DMP) for the E-TANDEM project. Specifically, the deliverable indicates how to handle the research data during and after the end of the project, providing practices and guidelines to indicate which data will be generated, processed and collected, to what extent this data will be publicly available, and how data will be curated and preserved (including after the end of the project). As it is a living document, the document will be adjusted, if/when needed, throughout the project's implementation.

1.1 E-TANDEM summary

E-TANDEM will develop and validate, at bench scale, the first direct process for the selective production of carbon-neutral higher oxygenate e-fuels (HOEF) from CO₂, water, and renewable power. The HOEF is composed of, in a first realization, a mixture of higher aliphatic alcohols with more than four carbon atoms (C₅₊); and in a second realization, a mixture of symmetric and asymmetric higher (C₈₊) aliphatic ether derivatives of said higher alcohols. The HOEF is liquid at ambient conditions and expected to unite high cetane index with mild oxygenate character. These features shall contribute to solve limitations associated with state-of-the-art hydrocarbon and heavily oxygenated (methanol, DME) e-fuels, showing advantageous logistics, superior combustion, and favourable tailpipe emission reduction properties for the heavy-duty marine and road transport sectors. The e-fuel production concept to be demonstrated integrates a first CO₂/H₂O high-temperature co-electrolysis step with the downstream thermocatalytic conversion of the resulting e-syngas (H₂+CO) stream into higher oxygenates in a slurry-phase tandem Fischer-Tropsch/reductive olefin hydroformylation/alcohol dehydration process. Moreover, the project will benchmark the new HOEF and blends thereof against relevant heavy-duty and marine diesel fossil fuels and carry out process techno-economic and life-cycle analyses to assess the soundness of current fleet-compatibility, sustainability, and societal aspects of the newly proposed e-fuel and its production concept.

1.1.1 Document structure

This document first describes the Findable, Accessible, Interoperable, and Reusable (FAIR) principles for data and the connection to the E-TANDEM procedure, and how the project will implement the FAIR principles. In chapter 3, a summary of the data collected and monitored throughout the project is provided. Chapter 4 focuses on the allocation of resources and how the FAIR principles and data management will be used to implement the resources. This chapter includes data security and ethical aspects. In Chapter 5 Conclusions and Recommendations are given. The last chapters 6–8 are of administrative nature, such as Risks and interconnections, Deviations from Annex 1, and the necessary references to complete the DMP.

2 FAIR Data and E-TANDEM

According to EC's guidelines, publishing research data must follow the approach "as open as possible, as closed as necessary" [2]. The objective behind this EC's incentive is to maximise the impact of EU-funded research projects and accelerate the technological implementation of findings through the interaction of the data produced in several research projects. Nevertheless, ensuring usefulness of the produced research data beyond the objectives of the single project is a challenging task [3]. The concepts of Findable, Accessible, Interoperable and Reusable (FAIR) data raised to ensure [4]: 1) proper collection, annotation, and archival of project data; 2) long-term care of valuable digital assets, with the goal that they should be discovered and re-used for downstream investigations, either alone or in combination with newly generated data. Thus, the DPM must ensure high-quality digital publications that facilitate and simplify the process of discovery, evaluation, and reuse in downstream studies. The following sections describe how E-TANDEM adopts the principles of FAIR data.

2.1 Making data findable

With regards to research data, E-TANDEM partners are committed to making the data findable, accessible, interoperable, and re-usable (F-A-I-R) as much as possible. This is done by providing a rich set of metadata describing the dataset and a standard, unique and persistent identifier, in case of scientific publications preferably Digital Object Identifier (DOI). The research data and bibliographic metadata needed to validate the results presented in the deposited scientific publications will be stored in the respective area of the repository. All deliverables that are intended for public use will also be made available for download via the E-TANDEM website. Additionally, partners will consider the Open Research Europe platform for submitting manuscripts, depending on the research results. The basic metadata that will be defined for each dataset and or document are the following:

- Title
- Description
- Creator and contributors
- Specification of the data identifier(s)
- Publisher (if relevant)
- Date of data collection
- Dates of document creation and last revision
- Language
- Keywords (defined by E-TANDEM partners)
- Acknowledgement of the EC funding of the E-TANDEM project

Ultimately, the Dublin Core Metadata Element Set (or a selection thereof) will be used to make the data findable.

2.2 Naming the data and files

To enable findability of data/research outputs the naming convention will be E-TANDEM-DATA_TYPE-DATE.EXTENSION.

In order to provide a quick identification and indexing, the consistent naming of all the data will function as metadata. Deliverables will for example be named and numbered in the following way:

E-TANDEM-DXX_TYPE-DATE.EXTENSION for an easy data identification within the project, and *E-TANDEM_<dnum>_<dname>_<sdis>_<ver>_<orgshortname>.pdf* to connect the type of data to the deliverable and organization when necessary.

- <dnum> represents the code of the deliverable
- <dname> the name of the deliverable as stated in the DoA
- <sdis> is filled with the acronyms related to dissemination level (e.g., sen = sensitive, pu = public)
- <ver> the version of the document.
- <orgshortname> organization that is performing the changes on the version

Versions 0.x will indicate that the document is still a draft. The first official document to be sent to the European Commission (EC) will be numbered as v1.0. Further revisions or new issues of a deliverable will make use of the following format: v1.x when drafting, and v2.0 when submitting and successively.

2.3 Making data openly accessible

E-TANDEM follows the EC principle to have data “as open as possible, as closed as necessary”. Therefore, data generated and collected throughout the project will, when marked as a Public Deliverable (PU), be stored in an open repository to be accessible to a wider community. Such deliverables will be in the Public Domain. Data which is made openly available will be on an Open Access website at the end of the project (www.e-tandem.eu). Should any specific instructions be required to facilitate data access, such as information relating to the nature and origin of the data, then these will be provided when the data is deposited. The data selected for open access will be made available without restrictions. The data relating to Deliverables which are classified at the consortium-level will be stored within the internal project platform (METT) with access restricted to the partners in the consortium. Additionally, deliverables that are considered confidential will be shared with stakeholders (if relevant) after signing an NDA. Moreover, a non-confidential version may be released after agreement from all partners to widen the sharing of knowledge and reach the highest impacts. E-TANDEM’s data strategy is in line with the aims for Open Science Policy under Horizon Europe.

2.4 Making data interoperable

In order to make data interoperable, all E-TANDEM (meta)data will use a formal, accessible, shared and broadly applicable language for knowledge representation. The interoperability of data/research outputs will be ensured by using standard processing software where possible (Windows Office, pdf, etc). Partners will arrange access to specialist software (e.g., modelling, data processing or software for physico-chemical methods) on a bilateral or project-wide basis, depending on the need and license possibilities. All data collected during the project will be exported to standard data formats (when possible) to facilitate interoperability. These may include .rtf, .txt and .docx for reports, and .csv and .xlsx for numerical datasets.



2.5 Increase data re-use (through clarifying licences)

At present, there is no standard regarding the licensing of data generated by E-TANDEM and it is likely that data licensing will be decided on a case-by-case basis. Where possible the most liberal approach will be used in accordance with Horizon Europe recommendations, with the exception of data that is of commercial confidential nature, which will necessarily be subject to more restricted licensing. If licensing the use of data arises, this will be noted in a future DMP update. Typically, data will remain usable (i.e., will not be outdated yet) for 24 months after the project completion.

3 Data Summary

Within the E-TANDEM project, a huge amount of data will be generated and collected across all work packages and activities, which might often be linked to each other. The purpose of the data collection within E-TANDEM is to validate the proper conception and development of the carbon-neutral higher oxygenate e-fuels (HOEF) from CO₂, water and renewable power, the hybrid catalysis production process thereof, and the cost, quality, sustainability, and environmental soundness of both the e-fuel and its production route.

3.1 Types of data and formats

Within E-TANDEM, two types of data have been identified: personal data and research data. Regarding the **personal data**, the compliance to the General Data Protection Regulation (GDPR) will be ensured). Any collection of personal data, as necessary for the operation, safety, and protection of use of, and access to partners' network, infrastructures, etc, will be used exclusively for the E-TANDEM purposes and no further transfer or disclosure shall occur unless required by law. E-TANDEM will be compliant with procedures for personal data processing, as set out in the EU Regulation EU/2016/679. With regards to **research data**, E-TANDEM partners are committed to making them FAIR as much as possible (see chapter 2 and 4).

In order to monitor and collect the data generated during the project, the project management team (CSIC and UNR) defined a dedicated template (Table 2-1), which will be discussed and updated during project General Assemblies (GA) and Work Package Leader Board (WPLB) Meetings. The template will be stored and accessible at the data-sharing platform Mett.

The file collects all necessary information as requested by the official template [1] (selection in the list below):

- Data Summary
 - Type and format
 - Origin of the data
 - Quantify and storage of the data set
- Fair data
 - Naming used
 - Identifiability / Mechanism used
 - Which data will be available and how
 - Possible restrictions
 - Timing for re-use
- Allocation of resources
 - Costs
 - Responsibilities

Table 2-1 enlists the main data types that are being generated or collected within E-TANDEM.

3.2 Data Utility

The collected data from E-TANDEM may be useful to a wide variety of stakeholders: scientific community working on electrolysis, catalysis, synthetic fuel production and characterization, life-cycle analyses, etc, decision makers (for the new designs), policy makers, expert in standardization and legislation sector, system operators, academia, potential end users, etc. In addition, the raw data sets may also be useful for other concurrent or future research projects.

Table 3.1 Standard table format

Data type/level	Typical data format
Raw research data	
Chromatograms, spectra, electrochemistry testing results, etc	ASCII, CSV, TXT, FID
Electron micrographs (SEM, FIB-SEM, TEM)	JPEG, TIFF, PNG
X-ray tomograms	XRM, TXRM
Scripts	HTML, RTF, MLX
Worked-out data	
Numerical, tabulated data	(Excel sheets) XLSX, Origin projects (OPJ)
1D and 2D chromatograms	JPEG, TIFF, PNG, EPS
ASPEN/HYSYS chemical process models & simulations; GABI LCA models; Mass & energy balances (all inputs and outputs included) – (Excel) Economic modelling	APW, HSC, GBMX, XLSX exported
Combination of data, image, and text data	DOCX, PPT
Process flow diagram and relevant info (chemical, T, P, requirements) – economic modelling and LCA	PPT, Excel
As part of reports/analyses	
Project internal communications, official project reports, literature review	DOCX, PDF, PPT

4 Resource allocation for FAIR data and data security

4.1 Allocation of resources

The responsibility for data management rests with the principal investigator at each of the project partners. The storage of project data (in for example the European Open Science Cloud (EOSC), an OpenAire, Registry of Open Access Repositories (ROAR) and the Directory of Open Access Repositories (OpenDOAR)) is free, however person months will be required to prepare the files for storage. Personnel costs for efforts for data handling are included in the project's budget. The potential values associated with the long-term preservation of data and the costs, including how data will be kept beyond the project, how long (at least 2 and up to 5 years) and how the costs will be met, will be discussed by the consortium in the coming months.

4.2 IPR Protection Strategy

Within E-TANDEM, efforts are focused on establishing a new e-fuel production route and new Intellectual Property (IP) for partners may be an outcome. Access to IP will be granted to ETANDEM partners if/where appropriate, in accordance with the Consortium Agreement. Open Science Practices and Intellectual Property Rights (IPR) protection activities will be managed under WP1/5. The IPR strategy of the project is to ensure that partners can fully exploit their market position through patent protection. The partners will sign a Consortium Agreement prior to the signature of the Grant Agreement with the EC, in which all the rules ensuring confidentiality (procedure for granting publication of the results), ownership regime, and IPR issues (background included and excluded, ownership, licensing terms, etc.) are detailed.

4.3 Data security

Project data will be stored in an open repository to make it accessible to a wider community based on the following classification level:

- **Public:** project data that can be published on E-TANDEM website/registries of scientific repositories and dissemination to a wider community
- **Confidential:** project data for which there is some obligation to third parties not to fully disclose; this data therefore will be disseminated to the public only upon the project partners agreement and in a restrictive way; this classification level includes deliverables categorized as "confidential" in terms of dissemination level and listed as such in the Grant Agreement
- **Internal:** project raw data used by consortium partners during the project, which is not available for publishing

Each consortium partner is responsible for proper storage, processing and sharing of the data created during their project activities. Each consortium partner is responsible for storage capacities of the data created during their project activities, for such data backup and recovery according to partner's organisation internal rules.

As indicated, the openly accessible data (i.e., the deliverables) and other project related documents will be stored by the project management team (CSIC and UNR) using the internal project platform (METT). UNR has a contract with METT detailing provisions for data security, continuity, and



availability. Further explanations on METT can be found in Deliverable 1.1 – Project Handbook. Responsibilities for data recovery and secure storage, when necessary, will be further upon agreement with the consortium, the Grant Agreement and the Consortium Agreement.

4.4 Other ethics and issues: use of external procedures for data management

Regarding the ethical issues, no issues have been identified or may arise from the research activities of E-TANDEM that could have an impact on data sharing. An appropriate Consortium Agreement has been elaborated to manage the ownership of results and access to key knowledge (Intellectual Properties Rights -IPR- and research data). The IPR strategy considers diverse commercialization channels including patents, copyrights, and trade secrets. IPR is handled in line with general EC policies regarding ownership, exploitation rights, confidentiality, commercial utilization of results, availability of the information, and deliverables (see chapter 4.2). The critical Intellectual Property (IP) required for the successful completion of the E-TANDEM project are either owned by the consortium members or the consortium has preliminary checked the freedom to operate, and IPR protection measures have been defined in the Consortium Agreement. A list of the most relevant patents by E-TANDEM partners can be found in Part B of E-TANDEM Grant Agreement No. 101083700. The background IP brought by partners will be available for use, remaining the property of the original owner. IP developed during the project will be owned by the partner(s) who develop it. The decision to publish IP will be made by the E-TANDEM consortium and the strategy for exploitation of the results will vary on a case-by-case basis, ranging from free access to licensing.



5 Conclusion and recommendations

This deliverable report presents the Initial Data Management Plan outlining the production, handling, preservation, and re-use of the (research) data generated in the E-TANDEM project. In addition, the methods that allow to make those data FAIR are demonstrated. The project consortium commits to abide to the provisions described in this report. This DMP is a living document, which will be updated throughout the lifetime of the project. A final DMP will be delivered at the end of the project (M42). For the project, it is vital the DMP remains current and relevant to ensure correct project data handling during and after the E-TANDEM project.



6 Risks and interconnections

No risks arisen related to this deliverable. This is a living document, which will be monitored and updated throughout the project lifetime.

6.1 Interconnections with other deliverables

This DMP describes data that will lead to results that might have possible exploitable value. Therefore, this DMP can also be seen as making clear, which partner(s) will produce particular data for possible future exploitation and intellectual property protection. The DMP serves as the initial guidance for these exploitations, and it will be updated throughout the project.



7 Deviations from Annex 1

There are no deviations from the description of this deliverable as given in Annex I of the Grant Agreement.



8 References

- [1] European Commission (2016) Guidelines on FAIR Data Management in Horizon 2020. Online available: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- [2] European Research Council (ERC), “Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020,” no. April, pp. 1–7, 2017.
- [3] European Commission, “Turning FAIR into reality: final report and action plan from the European Commission Expert Group on FAIR Data,” 2018.
- [4] M. D. Wilkinson et al., “The FAIR Guiding Principles for scientific data management and stewardship,” *Sci. Data*, vol. 3, no. 1, p. 160018, Dec. 2016, doi: 10.1038/sdata.2016.18.

9 Acknowledgement

The author(s) would like to thank the partners in the project for their valuable comments on previous drafts and for performing the review.

Project partners:

#	Partner short name	Partner Full Name
1	CSIC	AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS
2	MPG	MAX-PLANCK-GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV
3	DTU	DANMARKS TEKNISKE UNIVERSITET
4	OWI	OWI SCIENCE FOR FUELS GMBH
5	UNR	UNIRESEARCH BV
6	T4F	TEC4FUELS GmbH
7	AVL	AVL LIST GMBH
8	GF	GOODFUELS BV
9	UZ	SVEUCILISTE U ZAGREBU, FAKULTET STROJARSTVA I BRODOGRADNJE
10	UCT	UNIVERSITY OF CAPE TOWN
11	KAUST	KING ABDULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

Disclaimer/ Acknowledgment



Copyright ©, all rights reserved. This document or any part thereof may not be made public or disclosed, copied or otherwise reproduced or used in any form or by any means, without prior permission in writing from the E-TANDEM Consortium. Neither the E-TANDEM Consortium nor any of its members, their officers, employees or agents shall be liable or responsible, in negligence or otherwise, for any loss, damage or expense whatever sustained by any person as a result of the use, in any manner or form, of any knowledge, information or data contained in this document, or due to any inaccuracy, omission or error therein contained.

All Intellectual Property Rights, know-how and information provided by and/or arising from this document, such as designs, documentation, as well as preparatory material in that regard, is and shall remain the exclusive property of the E-TANDEM Consortium and any of its members or its licensors. Nothing contained in this document shall give, or shall be construed as giving, any right, title, ownership, interest, license, or any other right in or to any IP, know-how and information.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101083700. 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.