



Deliverable D6.4.

Data Management Plan (DMP): Initial Version

Date: 30/08/2024

AUTHOR: Lola Maldonado

CIC energiGUNE



This project has received funding from the European Union's Horizon Europe research and innovation programme through the European Innovation Council under the grant agreement No. 101130249.

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

DOCUMENT CONTROL SHEET

PROJECT INFORMATION

Project Acronym	WASTE2H2		
Project Full title	Plastic waste valorization to clean H ₂ and decarbonized chemicals through its catalytic deconstruction by novel ionic liquid-based catalytic systems		
Project Start date	01/03/2024	Project Ending Date	28/02/2028
Project Duration	48 months		
Funding call	HE-EIC-2023-PATHFINDEROPEN-01		
GA* n ^o	101130249		
Coordinator	CIC energigUNE		
Website	www.waste2-h2.eu		

*GA: Grant Agreement

DELIVERABLE INFORMATION

Deliverable n ^o - Title	D6.4 – Data Management Plan (DMP): Initial version	
Deliverable Due Date	31/08/2024	
Deliverable Lead Partner	CICe	
WP ¹ n ^o - Title	WP6 – Dissemination and Communication	
WP ¹ Leader	CIC energigUNE	
Task/s n ^o - Title	Task 6.1 – Dissemination and Data Management Task 6.2 – Communication	
Task Leader	Lola Maldonado – CICe	
Main Author	Lola Maldonado – CICe	21/08/2024
Other Authors	Eduardo J. Garcia Suarez - CICe Miriam Gutierrez – CICe	
Reviewer/s	Eduardo J. Garcia-Suarez – CICe	23/08/2024
Deliverable Type	Report <input checked="" type="checkbox"/> DMP ² <input type="checkbox"/> Demonstrator <input type="checkbox"/> DEC ³ <input type="checkbox"/>	
Dissemination Level	PU ⁴ <input checked="" type="checkbox"/> SEN ⁵ <input type="checkbox"/> Classified ⁶ <input type="checkbox"/>	
Date Approved by Coordinator	30/08/2024	
Status	Draft <input type="checkbox"/> Final <input checked="" type="checkbox"/>	

¹WP: Work Package; ²DMP: Data Management Plan; ³DEC: Websites, patent filings, videos, etc.; ⁴PU: Public, fully open; ⁵SEN: Sensitive, limited under the conditions of the GA;

⁶Classified R-UE/EU-R – EU RESTRICTED under the Commission Decision No2015/444

Classified C-UE/EU-C – EU CONFIDENTIAL under the Commission Decision No2015/444

Classified S-UE/EU-S – EU SECRET under the Commission Decision No2015/444

DELIVERABLE REVISION HISTORY

V.	Date	Author	Description of Change	Reviewed by
0	21/08/2024	Lola Maldonado – CICe	Initial version	Eduardo J. Garcia- Suarez – CICe 23/08/2024

List of Figures

Figure 3.1. Data storage systems.....	10
--	----

List of Tables

Table 3.1. Data description.....	11
Table 4.1. Main sources of project data.....	18
Table 4.2. Quality assurance questionnaire.....	20
Table 5.1. Other research outputs	22

List of Abbreviations

Abbreviation	Definition
ACS	American Chemical Society
API	Application Programming Interface
C	Confidential
CA	Consortium Agreement
CC0	Creative Commons Zero
CERN	
D	Deliverable
DD	Day
DMP	Data management plan
DOI	Digital Object Identifier
D&C	Dissemination and Communication
E&I	Exploitation & Innovation
EC	European Commission
EDX	Electron Dispersion Spectroscopy
EU	European Union
FAIR	Findable, Accessible, Interoperable and Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
HE	Horizon Europe
IGO	Intergovernmental Organization
IL	Ionic Liquid
IPR	Intellectual Property Rights
LCA	Life Cycle Analysis
LCC	Life Cycle Costing
M	Month
MD	Molecular Dynamics
MoM	Minutes of Meeting
MO-NP	Metal Oxide NanoParticles
MW	Microwave

Abbreviation	Definition
NGO	Non-Governmental Organizations
PID	Persistent Identifiers
PU	Public
R	Restricted
RP	Reporting Period
RSC	Royal Society of Chemistry
S	Secret
SEN	Sensitive
SOP	Standard Operation Procedure
T	Task
TEM	Transmission electron microscopy
TOF	Turn Over Frequency
TON	Turn Over Number
v	Version
WP	Work Package
XPS	X-Ray Photoelectron Spectroscopy
YYYY	Year

CONTENTS

1	EXECUTIVE SUMMARY	7
2	INTRODUCTION.....	8
3	DATA SUMMARY	9
3.1	Purpose of data collection and generation.....	9
3.2	Data storage system.....	9
3.3	Data description.....	11
3.4	Re-use of existing data	15
3.5	Data utility	15
4	FAIR DATA.....	17
4.1	Making data findable, including provisions for metadata.....	17
4.2	Making data accessible	18
4.2.1	Repository	18
4.2.2	Data.....	18
4.2.3	Metadata	19
4.3	Making data interoperable	19
4.4	Increase data re-use	20
5	OTHER RESEARCH OUTPUT.....	22
6	ALLOCATION RESOURCES	24
7	DATA SECURITY	25
7.1	Specific responsibilities regarding data protection.....	25
7.2	Long term preservation and curation	25
7.3	Confidentiality and integrity of data.....	26
8	ETHICS	27
8.1	Data sharing with confidential access.....	27
8.2	Archiving confidential information.....	27
8.3	Personal data.....	27
8.4	Handling of personal data.....	28
9	OTHER ISSUES.....	29
10	CONCLUSIONS	30

1 EXECUTIVE SUMMARY

This present plan comprises deliverable 6.4 - Data Management Plan (DMP): Initial Version. This deliverable is an integral part of Task 6.1 - Dissemination and Data Management and Task 6.2 - Communication, defined under the WP6 description, named Dissemination and Communication and led by CICE.

This plan is anticipated to encompass a set of protocols for overseeing the technical data produced throughout the WASTE2H2 project. The aim is to guarantee the accessibility and clarity of the generated data, aligning with the requirements of the "Open Research Data Pilot" guidelines. It is required to be submitted by the project's month (M) 6 of the project (end of August 2024). It will be updated in M30 (end of August 2026) with the D6.5 – Data Management Plan (DMP) RP2 update of DMP, and in M48 (end of February 2028) with the D6.6 – Data Management Plan (DMP): RP3 update of DMP, final version. This document will undergo continuous revisions as specified in the deliverables throughout the project lifecycle, ensuring that it remains up to date.

In general, WASTE2H2 will follow Horizon Europe's Open Access Publication policy by publishing scientific articles. In addition to well-known high-impact journals specific to their domain, WASTE2H2 will consider journals with the open peer review, such as Open Research Europe. Open Science and Open Innovation are integral to WASTE2H2's mission. The project will adhere to the Horizon Europe Access directives and embrace all avenues of Open Science. WASTE2H2 aims to share research early and openly except for data that could impact future commercialisation.

Overall, this document ensures a through coverage of all potential data processing needs, outlining the data management cycle for the data collected, processed and/or generated.

2 INTRODUCTION

Why is it essential to develop the WASTE2H2 Data Management Plan?

Projects chosen for funding through Horizon Europe must formulate a comprehensive Data Management Plan to ensure that their data and research outputs adhere to the principles of being Findable, Accessible, Interoperable, and Re-usable (FAIR).

The FAIR principles for effective data management in research projects include:

- Guidelines for managing research data during and after the project.
- The nature of the data collected, processed, and produced by the project.
- The methodologies and standards used.
- Plans for sharing or opening the data, including the methods of doing so.
- Strategies for curating and preserving the data.

This plan must be submitted as a deliverable in M6 of the project and periodically revised until the end of the project, M30 – August 2026 and M48 – February 2028) to reflect the project's evolving approach to data management.

The project coordinator, CICE, in their capacity as the leader of Work Package (WP) 6, will be responsible for implementing the DMP with the inputs from the rest of the partners.

3 DATA SUMMARY

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

What types and formats of data will the project generate or re-use?

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

What is the expected size of the data that you intend to generate or re-use?

What is the origin/provenance of the data, either generated or re-used?

To whom might your data be useful ('data utility'), outside your project?

3.1 Purpose of data collection and generation

The primary aim of WASTE2H2 is to pioneer a new methodology that combines innovative Ionic Liquid-based catalytic systems with microwave irradiation. This method aims to selectively produce highly pure hydrogen and valuable decarbonised chemicals, such as solid carbon, from plastic waste. WASTE2H2 endeavours to set the stage for a groundbreaking technology that enables the simultaneous production of pristine H₂ and carbon nanostructures from plastic waste, laying a foundation for innovative waste management solutions.

Consequently, the research conducted within WASTE2H2 will yield valuable data concerning the various components and the technology itself.

Each result generated in WASTE2H2 project will undergo assessment to determine whether it warrants protection for exploitation or if it's suitable for dissemination. This will be carried out as part of the ongoing monitoring and reporting and will be overseen by an Exploitation & Innovation (E&I) Manager as part of the Task 5.3 – Innovation, exploitation and IPR management, a role held by Dr. Cristina Domínguez from CICE. WASTE2H2 project data will be published as open data in accordance with the FAIR principle, if it is not subject to any commercial exploitation or access restrictions. These open data can be utilised by third parties to generate new outcomes. As an impression of data created in WASTE2H2, the following outputs are carried out: catalytic tests results (FAU, CICE, i-CLeHS), material characterization results (CICE), synthesis routes parameters (FAU, CICE), digital models (SENER ING, SENER MOB) and LCA/LCC inventory (FAU).

The mechanism for managing and providing accessibility to these results will be facilitated through platforms like the European Open Science Cloud (ZENODO), adhering to the FAIR Sharing Initiative guidelines and licensed under Creative Commons.

Overall, the purpose of this Data Management Plan is to comprehensively outline the key components of the data management policy that the WASTE2H2 consortium will adhere to for all data and datasets generated during the project.

3.2 Data storage system

Data storage technologies play a vital role in securing valuable information and shielding it from potential risks such as data loss, cyber threats, and system crashes.

There are many widely used ways for storing data, ranging from different electronic forms to conventional paper-based approaches.

In the context of this project, specific systems have been chosen to fulfil the diverse needs of information management and dissemination. These systems include:

- **Teams:** A project folder has been established on the Teams platform to streamline communication and collaboration among consortium members during the project. Within this folder, a wide array of resources is available, ranging from contact lists and templates to meeting minutes, agendas, presentations, deliverables, and the Technical Annex of the Grant Agreement (GA). These materials serve as essential references for project partners, ensuring seamless coordination and information exchange throughout the project lifecycle.
- **Zenodo:** As another integral component of the project infrastructure, Zenodo provides a reliable repository for storing and sharing research outputs, datasets, and other scholarly materials. This platform enhances the accessibility and visibility of project outcomes, fostering greater transparency and knowledge dissemination within the scientific community.
- **Website:** In addition to the platforms, a designated project website ([waste2H2 Project \(waste2-h2.eu\)](https://waste2-h2.eu)) has been established to extend the project's reach beyond its immediate stakeholders. This website serves as a centralised hub for accessing publicly available deliverables, project updates, and relevant news. By leveraging the website as an alternative communication channel, the project aims to engage with a broader audience and foster ongoing dialogue with the wider community.

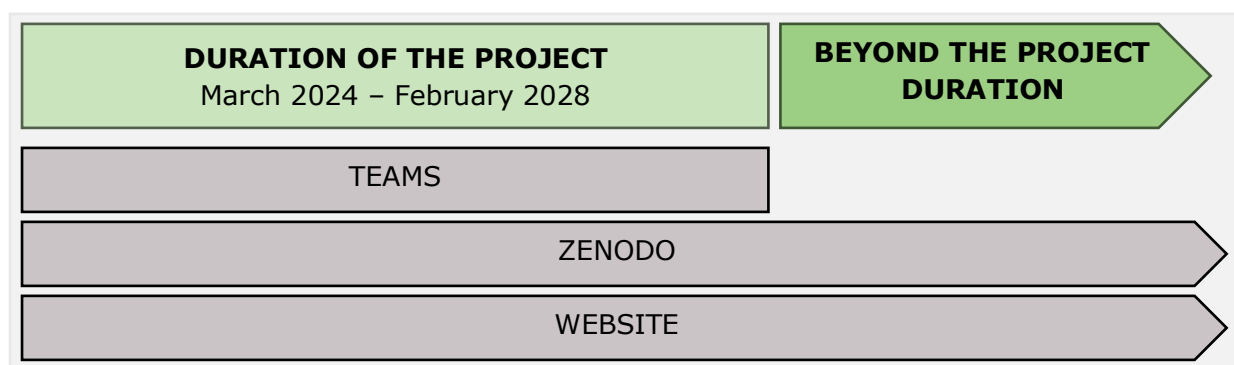


Figure 3.1. Data storage systems

Through the strategic integration of these storage and dissemination platforms, WASTE2H2 project endeavours to uphold the integrity of its data while promoting transparency, collaboration, and knowledge sharing among all stakeholders involved.

3.3 Data description

The following table represents the main data generated by the project:

Table 3.1. Data description

WP	Data type	Purpose	Data production methods	Partner Responsible	Data format/size	Useful for which stakeholder
WP2	<p>Data regarding:</p> <p>i) MD simulations of ILs, analysis of MD trajectories and material selection, including atomic coordinates, topologies, force field parameters, simulation settings, trajectory coordinates, velocities and forces, etc.</p> <p>ii) Screening of commercial ILs and development of novel ILs fulfilling required features.</p> <p>Data regarding:</p> <p>i) Main dielectric parameters of pre-selected ILs versus temperature and frequency;</p> <p>ii) Main dielectric parameters of predicted ILs versus temperature and frequency.</p>	<p>The input data is descriptive of the system properties and simulation algorithms/parameters to be used to model the studied system. The output data will be used to interpret the simulation output, extract patterns and observables to be compared with experimental data to gain further insight into the molecular mechanisms of the studied system.</p> <p>The data will be used for computationally screen ILs for specific features, select the most suitable IL from commercially available options, and design and synthesize custom ILs as needed. These efforts are geared towards optimising the selection and design of ILs for various applications,</p>	<p>Data will be produced via classical molecular simulation tools such as GROMACS.</p> <p>All participants and compiled in the project's shared Teams' SharePoint.</p> <p>Data will be produced using coaxial probe method at 25 °C – 200 °C over a frequency range of 10 MHz – 10 GHz.</p>	<p>CICe, IOL, FAU, i-CLeHS</p>	<p>Geometry structures in XYZ, PDB, MOL, GRO.</p> <p>Output files format is typically in CSV or plain text format. GitHub/CONDA/IP repositories will be in their respective formats.</p> <p>Data interpretation will be in either vectorial or rasterized formats (PDF, PS, SVG, PNG, TIFF, JPEG).</p> <p>MD simulations output: up to 200 k particles is in the order of Gigabytes.</p>	<p>WASTE2H2 consortium, Scientific Community, Industry</p>

WP	Data type	Purpose	Data production methods	Partner Responsible	Data format/size	Useful for which stakeholder
		fostering innovation in chemical synthesis and process optimisation. The data obtained for the dielectric parameters will be used to predict the ILs to be selected via a structure-property activity to enhance their performance under MW-irradiation.			Dielectric parameters are generally described and interpreted in PDF, PNG, TIFF or JPEG format. Size is negligible	
WP3	<p>Data regarding:</p> <ul style="list-style-type: none"> i) Properties of MO-NPs. ii) Synthesis and characterisation of MO-NP. iii) Preparation, characterisation, and stability studies of the novel IL-based catalytic systems. <p>Data regarding:</p> <ul style="list-style-type: none"> i) Main dielectric parameters of pre-selected MO-NPs versus temperature and frequency; ii) Main dielectric parameters of predicted MO-NPs versus 	<p>The data will be used to establish the criteria for the nature, morphology, and size of MO-NPs to maximise the MW effect, select, prepare, and characterise MO-NPs, accordingly, prepare IL-based catalytic systems, and study their thermal and chemical stability. Through these steps, it is aimed to unlock the potential of MO-NPs and ILs in catalytic processes, advancing our understanding of their stability and efficacy for diverse applications.</p>	<p>All participants and compiled in the project's shared Teams' SharePoint. Data regarding the heating profiles during the synthesis of MO nanoparticles will be produced to enable reproducibility of the synthesis; the particle shape and size will be assessed via Scanning and Transmission Electron Microscopy, stoichiometric composition via Electron Dispersion Spectroscopy (EDX). Relationship between MO-NPs and MW to</p>	FAU, i-CLeHS, CICE	<p>text format for the heating profiles, images in format .jpeg, .png, .TIFF a total output of 600 Megabytes. Dielectric parameters are generally described and interpreted in PDF, PNG, TIFF or JPEG format. Size is negligible.</p>	WASTE2H2 consortium, Scientific Community, Industry

WP	Data type	Purpose	Data production methods	Partner Responsible	Data format/size	Useful for which stakeholder
	temperature and frequency; i)Main dielectric parameters of pre-selected IL-based catalytic systems versus temperature and frequency; ii)Main dielectric parameters of pre-selected IL-based catalytic systems versus temperature and frequency.		enhance their performance will be done using an equipment analogue than coaxial probe (to be determined). Relationship between IL-based catalytic systems and MW to enhance their performance will be done using an equipment analogue than coaxial probe (to be determined).			
WP4	Data regarding: i) Results of the IL based catalytic systems under MW activation; ii) Product analysis (gas, solid and liquid); iii) Effects of the MW irradiation on the clean H2. iv) Continuous flow project solution.	The data will help to optimise reactions in semi-batch under MW irradiation, select the most efficient IL-based catalytic system, determine the best plastic substrate, analyse co-generated carbon for valorisation, understand MW irradiation's impact on conversion and selectivity, and test continuous-flow plastic waste deconstruction with the chosen catalytic system. It is aimed to	All participants and compiled in the project's shared Teams' SharePoint. Data will be produced using batch-microwave reactor and continuous flow-microwave reactor. Catalytic performance (TON, TOF, selectivity...) will be reported using conventional method. Quantitative and qualitative analysis will be obtained using	FAU, i-CLeHS, CICE, IOL, SENER ING	.doc or .xls. JPEG, .JPG, .PNG, .TIFF, AIFF, WAVE, .MP3, .MP4... Data from MW activation, catalytic performance and analysis are generally described and interpreted in PDF, PNG, TIFF or JPEG format. Size is negligible.	WASTE2H2 consortium, Scientific Community, Industry

WP	Data type	Purpose	Data production methods	Partner Responsible	Data format/size	Useful for which stakeholder
		advance understanding and application of MW-driven catalytic processes, contributing to sustainable waste management and chemical synthesis.	microGC for the gas, GC for the liquid and TEM, XPS, X-Ray for solid.			
WP5	Data regarding: i) Environmental and economic analysis. ii) Flexibility and accuracy to the new parameters. iii) Exploitation. iv) IPR management.	The data will help to develop a digital model that enhances flexibility and accuracy by adapting to new parameter values, conduct preliminary environmental and economic analyses through LCA and LCC, manage intellectual property rights (IPR), and identify opportunities for result exploitation. Through these objectives, it is aimed to streamline processes, optimise resource allocation, protect intellectual assets, and capitalise on opportunities for sustainable growth and innovation.	All participants and compiled in the project's shared Teams' SharePoint. LCA & LCC tool	All partners	.ppt, .doc files. Up to 100 MB CSV/XLSX files. Up to 20 MB PNG/JPG files. Up to 100 MB PDF files. Up to 50 MB	WASTE2H2 consortium, general public, Scientific Community, Industry

3.4 Re-use of existing data

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

What types and formats of data will the project generate or re-use?

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

What is the expected size of the data that you intend to generate or re-use?

What is the origin/provenance of the data, either generated or re-used?

The use of already existing data from research work and other European initiatives is strongly promoted. This section will undergo revisions in the later phases of the project and will be incorporated into the subsequent iterations of the Data Management Plan either in M30 or M48.

CICe will re-use data mainly focusing on literature data, existing synthetic and characterization routes, and databases for comparative purposes, including .pdf and .doc files.

FAU does not foresee re-use data at this stage. In any case, this information will be revised in the following update of this deliverable.

i-CLeHS will re-use literature data obtained through google scholar, or from the traditional publishers Elsevier, ACS, RSC, etc.

IOL will re-use already existing data in the form of literature and existing SOPs to produce standard ILs. Those are PDF/DOC files, used for the generation, synthesis and purification of the ILs needed within WP2 and WP3.

SENER will not re-use existing data, as their main part, the digital model, will use data generated during the project itself, and will not be based on previous models.

3.5 Data utility

To whom might your data be useful ('data utility'), outside your project?

The data generated (excluding any confidential information or document) is intended for reuse and utilisation beyond the project's duration as shown previously in Figure 3.1. The strategy for sharing this data is crafted to accommodate the needs of different stakeholders.

- **Project Consortium:** The data can be invaluable for consortium members for further research, analysis, and potential commercialisation opportunities. It can aid in informing future projects, collaborations, and decision-making within the consortium.
- **Industry:** Industries related to the project's domain can benefit from the data by leveraging it for product development, market analysis, and innovation. It can provide insights into consumer behaviour, market trends, and technological advancements, thereby enhancing competitiveness and driving economic growth.
- **Scientific Community:** Researchers and academics can use the data for advancing knowledge in the field, conducting further studies, and validating findings. It can contribute to the scientific discourse, support peer-reviewed publications, and facilitate interdisciplinary collaborations.

- **Society:** The data can have broader societal implications by enabling informed decision-making, fostering public awareness, and promoting transparency. It may be utilised by NGOs, advocacy groups, and educators to address social issues, advocate for policy changes, and enhance public understanding of relevant topics.
- **Policy Makers:** They can use the data to formulate evidence-based policies, regulations, and strategies. It can support informed decision-making, policy evaluation, and implementation of initiatives aimed at addressing societal challenges, promoting sustainability, and fostering innovation.
- **European Commission (EC):** The data can be aligned with the EC's objectives related to research, innovation, and societal challenges. It can contribute to EU-wide initiatives, frameworks, and policies aimed at promoting research excellence, economic growth, and societal well-being, thereby enhancing the EU's competitiveness and global influence.

Moreover, the data generated by the consortium will be categorised based mainly on the accessibility:

- **Public (PU):** fully open.
- **Sensitive (SEN):** limited under the conditions of the GA.
- **Classified:**
 - Classified R-UE/EU-R – EU RESTRICTED under the Commission Decision No2015/444.
 - Classified C-UE/EU-C – EU CONFIDENTIAL under the Commission Decision No2015/444.
 - Classified S-UE/EU-S – EU SECRET under the Commission Decision No2015/444.

4 FAIR DATA

FAIR data are data which meet the principles of Findability, Accessibility, Interoperability, and Re-usability.

- **Findable:** data that can be easily located and identified by both humans and computer systems.
- **Accessible:** means that the data can be storage in a way that allows to be accessed by both users and systems.
- **Interoperable:** data that can be easily integrated and used with other data.
- **Re-usable:** data that can be used for multiple purposes and by different users with no significant modification.

In accordance with the FAIR principles, Zenodo follows the standard of issuing a Digital Object Identifier (DOI) for all deposited datasets, as well as any related materials. This tool allows researchers to share and preserve their research output, such as datasets, publications, software, and presentations.

4.1 Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Will metadata be offered in such a way that it can be harvested and indexed?

The concept of findable refers to the viability of the information that it is located by other users, in other words, a good practice of organising and structuring data that can be easily located and accessed. Every time documents are uploaded in WASTE2H2 platforms, the following rules must be strictly followed as defined in D1.1 Project Management Plan:

File Naming Conventions: it is fundamental to implement a consistent and descriptive file naming conventions for uploading to Zenodo. This naming will be defined with the mandatory parts, and each of these elements will be separated by an underscore “_”.

- As a rule, files will be named as YYYY.MM.DD_WASTE2H2_Title, facilitating an easy track filtering it by date
- Final deliverables will be coded according to their title in the GA, e.g., WASTE2H2_DX.Y_Deliverable name

For the metadata, it is necessary to include, according to GA, the datasets with their description, date of deposit, author(s), venue, and embargo; Horizon Europe or Euratom funding; the grant project name, acronym, and number; licensing terms; persistent identifiers for the dataset; the authors involved in the action and, if possible, for their organizations and the grant; and, where applicable, persistent identifiers for related publications and other research outputs.

Use of persistent and unique identifiers. Persistent identifiers (PID) will be used to facilitate link between related information.

Searchable keywords. Descriptive tags will be added to files to make them easily searchable.

File versioning. All the documents will be clearly defined with sequential numbers (v0, v1, v2, ...) to indicate the chronological order of changes.

4.2 Making data accessible

4.2.1 Repository

Will the data be deposited in a trusted repository?

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

According to the FAIR principle, data being “Accessible” entails that it is stored in a manner enabling access by both users and systems. Zenodo, the platform chosen as the repository for WASTE2H2 as explained in section 3.2 of this document, offers a dependable and reputable platform for storing and disseminating research outputs and datasets.

4.2.2 Data

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Will the data be accessible through a free and standardized access protocol?

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

How will the identity of the person accessing the data be ascertained?

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

Zenodo, as part of OpenAIRE, offers a repository for researchers who lacks access to an institutional or thematic repository for depositing their publications and data. The following table provides a comprehensive overview of the primary sources of data used in this project.

Table 4.1. Main sources of project data

Data	Open Access location
Deliverables	HE results platforms, project website
Scientific publications	Zenodo, project website
D&C materials	Project website, social media (X, LinkedIn, YouTube)
Datasets	Zenodo

4.2.3 Metadata

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

According to Zenodo's code, which is itself an open source, all metadata is openly available under Creative Commons Zero (CC0) licence (except for email directions), and all open content is openly accessible through open APIs.

4.3 Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

Will your data include qualified references¹ to other data (e.g. other data from your project, or datasets from previous research)?

The WASTE2H2 project responsible for generating the data must also use appropriate methods, standards, procedures and established European metadata vocabularies to ensure its interoperability.

To maintain compatibility, all public documents generated during the project will be saved in .pdf, .txt, .rtf, .csv, .xlsx and .docx format. This practice will prevent issues with applications that read these documents.

For information that is published on the project website, all downloadable documents such as brochures and deliverables, and the videos uploaded in YouTube will adhere to .pdf and .html5 and .css web standards to ensure interoperability across different platforms and devices. Additionally, ensuring all multimedia files are in widely supported formats like MP4, MOV, WMV for videos and JPG or PNG for images will further enhance accessibility and compatibility.

The project's data that include qualified references to other datasets, both from the project and from previous research. Each deliverable that utilizes data will include detailed references to ensure traceability and context. This will involve citing relevant datasets, publications and any other resources that have contributed to the research.

¹ A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: <https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/>)

4.4 Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Will the provenance of the data be thoroughly documented using the appropriate standards?

Describe all relevant data quality assurance processes.

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

In WASTE2H2 project, ensuring increased data re-use is paramount for maximising the impact of the project results. To achieve this, several strategies will be implemented to facilitate the validation of data and encourage data re-use for third parties:

- Licensing of data. Zenodo will be used to upload the datasets; however, the license will be decided on a case-by-case basis and documentation that supports data analysis validation and reuse will be made available upon data deposit as accompanying materials. The method could be copyright, Creative Commons, open license, or other. However, when using Zenodo, all metadata is available under CC0. Besides, reused datasets facilitate provenance by documenting the sources from which they have aggregated content and the state of records at given time, the methods used to process them and workflows such as version control, history, linked metadata, etc. that they have in place to curate, share and preserve them. New datasets will include provenance information, where feasible directly in the metadata. In cases where data is reused, comprehensive documentation will be provided to support the integrity and reliability of the data. This documentation will be made accessible through Zenodo.
- Availability and embargo period.
- Re-use of data by third parties. Data will be available by third parties.
- Duration of data re-use. The data will be deposited in the Zenodo repository. This guarantees access and re-use of the information even after the end of the project.
- Quality assurance. A quality assurance questionnaire is developed to ensure that datasets can be submitted to the EC. It should be answered by all the responsible of the quality assurance of each project partner.

Table 4.2. Quality assurance questionnaire

Question	CICe	FAU	i-CLeHS	IOL	SENER
Is the data collection methodology clearly defined and followed?	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
Is the data accurately and consistently recorded?	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
Does the data comply with the ethical standards and regulations?	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
Are the data storage procedures secure and in line with GDPR	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No

Question	CICe	FAU	i-CLeHS	IOL	SENER
requirements?					
Is there a backup and recovery plan for the data?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Is the data properly anonymized or pseudonym if required?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Is the data accessible to all the relevant project partners?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Are the data formats and standard consistently applied across the project?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Has the data been reviewed for accuracy and completeness?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Are the metadata and documentation adequately provided and maintained?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Is there a clear data sharing policy in place?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Have all data quality issues been addressed and resolved?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>
Is the data ready for submission to the EC?	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>	<i>Yes/No</i>

This questionnaire aims to cover all critical aspects of data quality assurance to ensure that the data meets the standards required for submission to the EC. However, this questionnaire will be reviewed, answered and filled out each time a dataset is to be uploaded in Zenodo. Therefore, in the following versions of this deliverable, the update of the mentioned questionnaire for each dataset will be included.

5 OTHER RESEARCH OUTPUT

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

Table 5.1. Other research outputs

WP	Data type	Purpose	Data production methods	Partner Responsible	Data format/size	Useful for which stakeholder
WP2	Data regarding: i)GitHub/CONDA/PIP repositories for any ad-hoc developed code ii)Calculation protocols documented either in text format or with workflow managers, if applicable iii)Plots and graphs representative of the studied observables	The input data is descriptive of the system properties and simulation algorithms/parameters to be used to model the studied system. The output data will be used to interpret the simulation output, extract patterns and observables to be compared with experimental data to gain further insight into the molecular mechanisms of the studied system.	Data will be produced via classical molecular simulation tools such as GROMACS.	CICe	Geometry structures in XYZ, PDB, MOL, GRO. Output files format is typically in CSV or plain text format. GitHub/CONDA/PIP repositories will be in their respective formats. Data interpretation will be in either vectorial or rasterized formats (PDF, PS, SVG, PNG, TIFF, JPEG). MD simulations output: up to 200 k particles is in the order of Gigabytes.	WASTE2H2 consortium
WP5	i) Git repository	The input data for the digital model characterizes the system properties and the	Classical molecular simulations, as	Sener	JSON files. Up to 500 MB	WASTE2H2 consortium

WP	Data type	Purpose	Data production methods	Partner Responsible	Data format/size	Useful for which stakeholder
	ii) Calculation protocol documented in pdf format	algorithms/parameters used for simulating the studied system. Data provided will be used for performing system simulations as developed during the project.	machine learning methods (using scikit-learn and keras).		CSV/XLSX files. Up to 20 MB PNG/JPG files. Up to 20 MB PDF files. Up to 20 MB	

6 ALLOCATION RESOURCES

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

Who will be responsible for data management in your project?

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

CICe will be the responsible for data management as part of the WP6, designated under “Dissemination and Communication” name. CICe has allocated a part of the overall WP6 budget and person months to the activities related to the DMP. There are no specific direct costs or indirect cost related to data management within the project.

Long-term preservation of data will be managed by the coordinator as following:

- During the project, SharePoint (Teams) data will be maintained during the project duration and will be managed by the coordinator.
- During and after the project, various measures are expected to be put in place to effectively oversee, preserve, and guarantee data. These measures are the following:
 - Secure and Fortified Data Storage: Comprehensive security measures will be enacted to protect the projects’ data. Zenodo will be used as a secure repository, and will be implemented to prevent data loss, unauthorised access, or corruption.
 - Detailed Documentation: A thorough documentation process will catalogue the structure, format, content, and metadata of all generated datasets. This detailed documentation will provide a roadmap for future researchers beyond the project's duration, building up the usability and understanding of the data.
 - Continued Data Sharing and Dissemination: WASTE2H2's commitment to data sharing and collaboration will continue post-completion.
 - Sustainable Long-Term Preservation: Strategies for long-term preservation will be developed to ensure the lasting accessibility and integrity of the project's data. These strategies may include sticking to regular data backups, standardised file formats, and adopting emerging technologies to reach future-proof data preservation efforts. For an in-depth discussion on these long-term preservation strategies, refer to section 7.2

By implementing these measures, the project aims to ensure the ongoing availability and utility of its data for future research. Continuous refinements and updates to the DMP will be made to maintain transparency and adherence to best practices in data management after WASTE2H2 project’ completion.

7 DATA SECURITY

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

Will the data be safely stored in trusted repositories for long term preservation and curation?

7.1 Specific responsibilities regarding data protection

Project consortia shall cooperate to ensure with legal obligations arising under applicable data protection laws, including Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data and on the free movement of such data and relevant national data protection law applicable. This cooperation is within the scope of performing and administering of the Project and of the Consortium Agreement (CA) already signed by all partners.

The following recommendations are made to WASTE2H2 project's consortium:

- Develop and implement a comprehensive data security plan.
- Educate employees on data security best practices.
- Use encryption and other security technologies.
- Regularly monitor and test security measures.
- Comply with relevant regulations and standards.

Zenodo is hosted at CERN, which enjoys a special legal status as an Intergovernmental Organization (IGO), offering data protection standards comparable to the EU's General Data Protection Regulation (GDPR). Specially on Zenodo, the following activities are considered:

- Reviewing and updating the Privacy Policy to provide specific information about the collected data, how it is used and when it is deleted.
- Assessing the collection of personal information.
- Evaluating and ensuring that any third-party tools and services are GDPR compatible.

7.2 Long term preservation and curation

One of the Zenodo and WASTE2H2 project's objectives is to guarantee the long-term preservation of all documents in digital format, in order that it can be read or reproduced:

- The content will be checked regularly to preserve its integrity, security, and durability.
- The content will be transformed into new formats when deemed necessary, based on the same security and durability criteria.
- It will provide, where possible, software emulations for access to formats that cannot be migrated.
- Regular backups will be provided, including data and metadata.
- Hardware and software updates, keeping the latest versions of the software used in the repository.

The consortium employs a federated approach, with each partner responsible for collecting all data, ensuring its security, and maintaining long-term preservation by using Zenodo as the trusted repository.

7.3 Confidentiality and integrity of data

The confidentiality and integrity of data are maintained through both partners' data access policies and the security measures of the storage system. WASTE2H2 project's partners enforce a strict policy on all employees, co-workers, and subcontractors with data access. This policy includes, but it is not limited to:

- Allowing copies on local devices only during data processing, with guaranteed erasure afterward.
- Extending access control policies to local copies.
- Implementing contractual clauses.
- Requiring agreement to terms and conditions before granting access.

Additionally, data stored in the Zenodo repository is protected against unauthorized access through the following security measures and functionalities.

- The data centre is located on CERN premises, with physical access restricted to a limited number of trained staff granted access based on their professional duties.
- Servers are managed according to CERN Security Baseline for Servers, restricting remote access to trained Zenodo staff, and ensuring the operating system and applications are updated with the latest security patches via the automatic configuration management system.
- The CERN Security Team runs host and network-based intrusion detection systems, monitoring traffic flow, patterns, and contents into and out of CERN networks to detect attacks.
- Zenodo stores user passwords using cryptographic hashing algorithms.

8 ETHICS

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

8.1 Data sharing with confidential access

Every time WASTE2H2 project's partners deal with confidential data, they must ensure that they are complying with the non-disclosure policy detailed in Section 10 "Non-disclosure of information" of the CA. The CA is already signed by all partners. For a period of 5 years after the end of the project, they will undertake in addition and without prejudice to any commitment on non-disclosure under the GA.

- Not to use confidential information for any purpose other than the one for which it was disclosed.
- Not to disclose confidential information to any third party without prior written consent from the Disclosing Party.
- To ensure the internal distribution of confidential information within a partner's organisation occurs strictly on a need-to-know basis.
- To return to the disclosing party, or destroy, all confidential information disclosed to the partners upon request, including all copies, and to delete machine readable information to extent practically possible.
- The partners may retain a copy of the confidential information only if required by applicable laws and regulations or for proof of ongoing obligations if they comply with the confidentiality obligations herein with respect to such copy.

8.2 Archiving confidential information

As stated in the guidelines outline in Section 10 of the GA, partners organisations tasked with collecting sensitive data will securely store it on internal servers, employing stringent security measures.

To ensure the confidentiality and integration of the information of these files, the coordinator will regularly update the access to the information in the Teams folder.

8.3 Personal data

The WASTE2H2 project's consortium partners are committed to upholding ethical and research integrity principles throughout the project. They will conduct all project-related activities in accordance with ethical principles outlined in documents such as in the European Textbook on Ethics in Research (EU, 2010), in the document Ethics for researchers (EU, 2013), and in the European Code of Conduct for Research Integrity (All European Academies, 2017), as well as applicable international, EU, and national laws. Additionally, the project will adhere to the European Human Rights Convention, particularly regarding privacy and autonomy.

Personal data will not be collected during the project activities, except for the contact details of stakeholders interested in receiving information about WASTE2H2 project's activities through specialised events. Their Personal data (contact list) will be handled and

processed in accordance with National and European Regulation, specifically the GDPR (EU2016/679). All consortium partners will comply with this requirement.

8.4 Handling of personal data

Personal data (or datasets that contain personal data) gathered by a consortium member will remain under the care of that member and will not be distributed to any other partner or any external partner. If a dataset contains personal data, the consortium parent responsible for creating the dataset must remove any personal data from the original dataset and create a separate storage for the personal data (e.g. identifying information) for referencing purposes. This referencing is necessary to identify the data of a particular individual should that individual wish to withdraw their data from use by WASTE2H2 project.

9 OTHER ISSUES

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

WASTE2H2 project does not initially plan to use of other national, funder, sectorial or departmental procedures for data management. Therefore, no additional procedures will be listed or described in this context.

10 CONCLUSIONS

This DMP delineates the protocols governing the acquisition, organisation, and processing of data utilized throughout the WASTE2H2 initiative. Serving as a comprehensive handbook, this document is indispensable for all project partners to proficiently manage their data.

Inclusivity is ensured through the integration of feedback from project beneficiaries into this document, which will be accessible to all upon finalisation, serving as a point of reference.

This DMP also addresses the dissemination of data, ensuring that it is shared and made accessible to project partners and other stakeholders in accordance with open access policies and relevant regulations.

Two different updates to this deliverable will occur at M30 and M48, being this last update the final version of the DMP. Nonetheless, any pertinent modifications within this timeframe will be documented and shared with partners as necessary.



waste2H2

