Tech UPGRADE

D1.2 Data management plan

Funded by the European Union

This project has received funding from European Union's Horizon Europe's Research and Innovation Program under grant agreement No. 101103966. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

Deliverable 1.3

Actual Submission Date:	**31/10/2023**
Produced by:	**Technical University of Denmark (DTU)**

TechUPGRADE

techupgrade.eu

HORIZON-CL5-2022-D4-01

Thermochemical Heat Recovery and Upgrade for Industrial Processes

Grant Agreement no.: 101103966

Start date of project: 1 May 2023 - Duration: 48 month

DELIVERABLE FACTSHEET

Deliverable D1.3	
Nature of the Deliverable:	R - Document, report
Due date of the Deliverable:	M6 – 31/10-2023
Actual Submission Date:	M6 – 31/10-2023
Produced by:	DTU: Ahmad Arabkoohsar
Contributors:	DTU: Hamid Reza Rahbari
Work Package Leader	DTU: Ahmad Arabkoohsar
Reviewed by:	DTU: Ahmad Arabkoohsar

Dissemination level	
X	PU = Public
	PP = Restricted to other programme participants (including the EC)
	RE = Restricted to a group of the consortium (including the EC)
	CO = Confidential, only members of the consortium (including the EC)

Contents

Tables

1 Summary

This document describes the Data Management Plan (DMP) for the TechUPGRADE project. The DMP sets out the consortium's approach to managing project data and provides links to other sources such as EC's official summary of the General Data Protection Regulation (GDPR) [1], Grant Agreement [2], and Consortium Agreement [3].

2 Introduction

The DMP outlines the strategy for managing data collected and/or generated by the TechUPGRADE project. This document, among others, defines a common understanding of project data, establishes mechanisms for exchanging and storing data, and describes how and what will be publicly available. The document is in line with Horizon Europe FAIR data management guidelines **[4]**, where FAIR stands for "findable, accessible, interoperable and reusable". In order to ensure that the project's data are FAIR for both manual and automatic use, follow these steps:

•**Findable** – The project's data should be deposited in a trustworthy repository. When added to the repository, data and supplementary documents receive a persistent identifier (PID) such as DOI, and are described by rich metadata (such as subject terms, provenance information, etc.).

•**Accessible** – Where appropriate, access to data should be provided under certain conditions or restrictions, with (meta)data being retrieved through PIDs assigned by a repository.

•**Interoperable** – Data interoperability is understood as the ability to connect data with other datasets, applications, and workflows. This can be achieved: (i) by using well-known and preferably open-data format and software whenever possible, (ii) by applying relevant standards for metadata, and (iii) by employing community-accepted schemes, keywords, controlled vocabularies, thesauri, or ontologies where possible.

•**Reusable** – Data should be well-described with metadata and provenance information, e.g., on how, why, and by whom the data have been created and processed. Further, an accessible data usage license should be available for others to know what kinds of reuse are permitted.

The TechUPGRADE project aims to comply with the GDPR **[1]** governing the EU citizen's personal data by other parties. Overall, there are 7 funding principles of GDPR:

•**Lawfulness, fairness, and transparency** – Processing of personal data is a must and cannot be conducted in another way with lesser impact on the data subjects. Clear information about the processing of personal data has to be provided to the data subjects, taking into account the purpose of personal data processing and information on whether these data will be used by other parties.

•**Purpose limitation** – Personal data may only be processed for specific and clearly defined purposes.

•**Data minimization** – Personal data may only be processed to the extent necessary for the purposes for which they are processed.

•**Accuracy** – Personal data must be accurate and up-to-date.

•**Storage limitation** – Personal data must be deleted or anonymized as soon as identification of data subjects is no longer required.

•**Integrity and confidentiality** – Personal data must be processed in a manner that ensures adequate security, including protection against unlawful processing and accidental loss, destruction, and damage. To protect personal data, an effective system of anonymization or pseudonymization should be conducted.

•**Accountability** – All consortium partners are responsible for compliance with the GDPR principles. This requires from each organization an implementation of appropriate technical and organizational measures and the ability to demonstrate, upon request, what has been done and its effectiveness.

This deliverable is a living document and might be updated throughout the project.

3 Project Data

3.1 Type of data

During the TechUPGRADE project implementation, both experimental and numerical data will be generated. These results will be opened to the public, with only a few exceptions due to data confidentiality, i.e., D1.1 – Project management plan (D1.3 Quality Management Plan), D2.1 Design selection of the heat upgrade reactors, D5.3 Software as the digital twin of the system, D6.1 Plant design of the pilot units, D8.7 Exploitation and sustainability plan. A complete overview of the data produced during the project lifetime, including their type and confidentiality status, can be found in Annex I of the Grant Agreement [2]. An important aspect of making data findable is to attribute a suitable naming convention to project data and documentation. The naming and versioning convention of datasets follows a structure agreed on by the consortium partners (DTU – Project Coordinator, COWI, GRL, UT, DLR, TUW, RISE, MGS, ABS, KU, IED, NTUA, TMEC, QTL), as addressed in D1.3 – Quality management plan, and referenced below.

For project deliverables, the following naming convention applies:

- Draft: <Deliverable Number><Deliverable Title><Document Version><Author> (e.g., D1.3 Quality management plan v0.1 MM)
- Final to the CINEA: <Deliverable Number><Deliverable Title><Document Version> (e.g., D1.3 Quality management plan v1.0)

The versioning management of project deliverables has the following steps. Before the review process, the author(s) will assign the document version to v0.1, which will be further incremented to v0.2 only by the author. In turn, the reviewers will apply the next available number behind the version number v0.1.X, where X is the next available number. When final, the document version will be changed to v1.0 by the Coordinator. If rejected by the granting authority, the document versioning will start with v1.1 and will be submitted as v2.0. The other data sets should be named as follows: <WP Number><Title><Date>, where “title” is a meaningful short description of the file (e.g., project meeting/workshop/etc.), and “date” is YYYYMMMDD when the activity took place. For example, the Kick-off meeting of TechUPGRADE project 20230523.

3.2 Data formats

To prevent loss of access to files and to enhance the long-term interpretability of (meta)data, data will be saved in the form of easily accessible and low-volume files.

Recommended file formats are indicated in Table 1. For long-term storage, files will be converted to open file formats whenever possible [5].

Table 1: Recommended file formats for the TechUPGRADE project

Item	Format
Text	TXT, XML, HTML, PDF, DOC, DOCX
Presentation	PPTX, PPT, PPF
Spreadsheets	MSE, CSV, Tab-delimited values, PDF, XLS, XLSX
Images	JPEG, TIFF, PNG, SVG
Numerical	NetCDF, CSV, JSON, TXT
Video	AVI, MP4
Audio	WAVE, MP3
Databases	Delimited Flat File w/DDL
Archives	ZIP, GZIP, 7Z, TAR

Each dataset and its associated documentation will be deposited in the dedicated channels in the Microsoft Teams (SharePoint) platform (MS Teams). To increase the interpretability of the data, README files will be created along with the project documentation. The README file provides information about the dataset and helps ensure that the data can be correctly interpreted when shared or published. In general, for each dataset one README file should be created. The README document should be written as a plain text file and saved as README.txt or as README.pdf. The recommended content of a README file is presented below [5].

- Introductory information:
  - Dataset title,
  - For each file or group of similar files, a brief description of data it contains,
  - Explain the file naming convention, if applicable,
  - File format if not evident from the file name,
  - If the dataset is built from multiple files that are related to each other, the correlation between the files or a description of the file structure that holds them,
  - Contact information (for the users questions about the data files).
- Methodological information:
  - Description of the data collection or generation method, as well as the data processing methods if non-raw data is provided,
  - Any instrument-specific information required for understanding or interpretation the data,
  - Software (with the version number included) that is used to create, compress, analyze and/or required for reading the dataset, if applicable,
  - Standards and calibration information, if applicable,
  - Description of any quality assurance procedures conducted on the data,

  - Definitions of codes or symbols applied for noting or characterization of the dataset of poor quality, questionable nature, and outliners, which is of interests for the users.
- Data specific information:
  - Full names and definitions of column headings for tabular data, with abbreviations expounded,
  - Measurement units,
  - Definitions of codes or symbols applied for recording the missing data,
  - Specialized formats or abbreviations are used.
- Sharing and access information:
  - Licenses or restrictions imposed on the data, allow to define the terms of use for the datasets.

Openly accessible data can typically be accessed, used, reproduced and distributed free of charge to the user. Otherwise, dataset should be accompanied by a data license (e.g., CC0, CC-BY) that details the permission associated with the use of the dataset. The licenses for datasets will be stored in the dedicated channels in MS Teams, and will be referenced in the respective dataset's README file and all metadata files.

3.3 Data storage

Overall, for the TechUPGRADE project the following data storage manners will be used: (i) The Microsoft Teams (SharePoint) platform (MS Teams), (ii) the project's website, (iii) repositories like at DTU OneDrive or external repositories [5]. All datasets generated or collected within the TechUPGRADE project will be handled in MS Teams. This is to facilitate collaboration and ensure transparency among the project partners with the location and accessibility of the project's documentation. For easy identification of datasets, a series of (sub)channels is created. Access to MS Teams will be managed by the Project Manager. An open-data will be also available on the project website, which will be used for internal and external communication needs and as an additional repository for deliverables. Selected datasets underlying scientific publications will be made publicly available using the public repository DTU OneDrive Datasets then get a DOI/persistent identifier and will be citable in publications. The project's data will be archived in DTU Data [6]. As for the data storage in portable devices (e.g. external hard drive, USB stick) the data storage is intended as short-term and backed up regularly on a non-portable storage. If portable devices have to be used for confidential data, devices have to be encrypted. Data, including that of a confidential nature, on the portable devices, will be deleted as soon as possible and no later than ending the activity for which the data will be needed.

3.4 Data sharing

The project's outputs will be made available to academia, industry, public authorities, and policymakers at the transnational level through, e.g., the project's website, OA publications in scientific journals (Joule; Nature Energy; Applied Energy; Energy Conversion and Management; Energy; Applied Thermal Engineering; Advanced Sustainable Systems; J. of Cleaner Production; Solar Energy Advances; Solar Energy; IEEE Access; Issues of Chemistry and

Chemical Technologies; Open Research Europe), and conferences and other dedicated events (nt. Conf. on Efficiency, Cost, Optimization, Simulation and Environmental Impact; IRES; IMPRES; ENERSTOCK; IEA SHC; ICCMA – Int. Conf. on Control, Mechatronics & Automation; ISES Solar world congress; EUROSUN; 4DH Smart Energy) to maximize the project's results visibility.

In this regard, open science practices based on the principle "as open as possible, as closed as necessary" will be applied by partners of the TechUPGRADE project. This refers to sharing the project's results widely and transparently while respecting the confidentiality of this data when required for legitimate obligations. These practices include, e.g., prompt and open access to the project's outputs through publishing in high-quality OA journals, preferably with a license for reuse (e.g., CC0, CC-BY), depositing data in trusted repositories, such as DTU data, and involving relevant knowledge actors in co-creating the content (see GA, Annex 1 [2] and Horizon Europe Programme Guide [7] for additional information). Further, project partners will apply their home institutions' systems for making data an open access, and simultaneously support other consortium partners in these activities.

Within the consortium, data will be shared with the use of MS Teams and email, with the former remaining a default platform for all project documents. Data with personal information will be shared through the channels that comply with the GDPR. For this purpose, e.g. SharePoint can be used to securely send digital data to others or receive data from others. For sensitive data, e.g., personal data or commercially interesting data, it is advised to use the encrypted option of SharePoint, with the key to decrypt these files sent separately.

3.5 Data archiving

In general, archiving datasets involves selecting and organizing files with data and related materials, preparing data documentation, and generating archival files, including encryption of sensitive data. For selecting and organizing files for preservation, attention should be given to the following:

- Ensuring that data and related materials are correctly chosen and concluded;
- Restricting data files to data alone and including, e.g., figures and analyses based on this data in separate files;
- Grouping data into smaller and larger files, not just files of one dominant size, considering that extensive files may exceed the software capacity. Some examples of data aggregation include type, location, period, measurement platform, investigator, method, or instrument.

Note that the beneficiaries should follow the archiving rules within their institutions.
As follows from GA, Data Sheet [2], after the final payment the beneficiaries are obligated to keep records and other supporting documents to prove the proper implementation of the action in line with the accepted standards in the respective field until the time limits indicated below:

- Confidentiality: 5 years
- Record keeping: 5 years
- Reviews: 2 years
- Audits: 2 years
- Extension of findings from other grants to this grant: 2 years
    - Impact evaluation: 5 years.

Data generated during the project will be stored at the partnering institutions' servers in compliance with applicable EU and national law regulations regarding data management and ethics. the Project Coordinator will securely retain the entire project dataset for a period of 10 years. This data will be stored on servers certified under ISO 27001 and NEN 7510 standards, specifically within the Faculty data storage and DTU facility, and will be managed in compliance with DTU's data policy. It is anticipated that the project will generate an amount of data that does not exceed 1 TB.

4 Dissemination and exploitation of the results

As follows from GA, Article 17 [2] and CA, Article 8 [3], the consortium will ensure adequate dissemination and exploitation of the project's activities. The Exploitation Committee (EC) supported by the Exploitation Manager will be also established to ensure proper attention to the importance of dissemination and exploitation of the project's outcomes and their visibility at the transnational level. For this purpose, the Dissemination and exploitation plan (D&E plan) will be written and reported in 4 common deliverables (D8.1, D8.3, D8.4, D8.7), with the latter describing, among others, the prospects of patenting and further exploitation of the generated outcomes. Also, the D&E plan will represent dissemination activities of the project, communication themes, and channels, target audiences by type and dissemination goals, and dissemination objectives in the short, medium, and long term. Overall, the following activities will be conducted as part of the dissemination and exploitation strategy:

- Publishing in scientific journals and specialized magazines (as Open Access), patents, and licenses;
- Participating in conferences, seminars, clustering events, international expositions and fairs;
- Organizing capacity-building activities like workshops, and boot camps;
- Sharing project results on popular repositories for the research data (e.g., DTU data, OneDrive Research Europe, Zenodo, OpenAir);
- Distribution of designed promotion material (e.g., newsletters, brochures, press releases, etc.);
- Integrating the project's results in the courses of the consortium partners.

The Exploitation Manager with the support of the Project Management will ensure a secure use of all GDPR-related data of the identified target audiences.

5 Allocation of resources

The data management task will be carried out by DTU under WP1 – Project Management and Coordination, although the consortium as a whole is required to comply with data management regulations as indicated in GA, CA, and DMP (in that order). Project partners included in their budgets the costs of transferring the project's data into FAIR data. Storing and preserving this data by DTU in dedicated internal and external archiving servers will be free of charge. However, repositories of specific institutions might not be free. In this case, project partners will use their budget for data storage and preservation, by the relevant institute's protocols. Overall, the project is expected to generate less than 1 TB of data.

6 Confidentiality and data protection

Although the project's results will generally not be confidential (13 of the 33 results are sensitive), the consortium will comply with applicable EU and national data protection laws (including authorization or notification requirements) about sensitive data, i.e., any data that reveals a subject's information. This data will be treated under dispositions of GA, Article 15 [2] and CA, Article 4 [3], and in line with GDPR [4]. Guiding principles for data protection and confidentiality can be found below.

- Project partners are responsible for compliance with GDPR, and doing the GDPR registration at their respective institutes.
- During the project's lifetime and 5 years after the final payment, all partners are obligated to maintain the confidentiality of any undisclosable information, with exceptions regulated by GA, Article 15.
- Overall, only sensitive data that is related to the project's activities and kept to an absolute minimum. This data will be verified for consistency and currency.

The consortium will ensure transparency in the collection of personal data. This will be realized, among others, by taking the necessary privacy and security measures in the project's communication media. In the case of collecting personal data as part of, e.g., surveys, questionnaires, interviews, workshops, trainings, etc., the following information will be communicated to the respondents:

- Type of information,
- How the data will be collected and processed,
- Whether, how, and for what purpose it will be disseminated,
- Whether and how it will be made an open access.

Overall, a data subject will have the possibility to request information on the type of sensitive data collected and to request its deletion to a reasonable extent. All personal data will be immediately anonymized/encrypted and deleted as soon as unused for the project.

7 Intellectual property rights

All obligations of project partners related to intellectual property rights (access rights to pre-existing know-how and knowledge necessary for the project execution, rules for dissemination, and use of own knowledge generated during the project) are settled in the CA [3], and further described in the GA [2]. Below are some key regulations.

- Overall, foreground knowledge will be royalty-free, whereas background knowledge will be available on predefined terms.
- Project partners must ensure mutual access to the background knowledge identified as needed for the project implementation – see GA, Articles 16, and is binding for the background knowledge identified in CA, Appendix 1. If a partner wishes to modify or withdraw its contribution from CA, Appendix 1, the consent of the General Assembly is required.
    - Granting the access rights not covered by GA or CA will be at the sole discretion of the owning partner, under conditions agreed between the owning- and receiving partner.

- Partner owns the generated results. Note that two or more partners can own results jointly under the settled conditions (e.g. joint ownership, or one owner but favorable licenses for the other parties, etc.), and by GA, Annex 5, and in CA, Article 8.
- Any knowledge produced during the project and claimed confidential by one of the project partners will be declared as such when disclosing it to the consortium or, if not marked as confidential – a written confidentiality statement will be provided within the shortest possible time after the disclosure of such information.
- Public data will be published after they are reviewed and approved by the grant authority. Before the publication, a written notice will be provided to the Coordinator. This procedure is also valid for data generated within the project and disseminated later (according to the post-project impact policy).
- Each partner will implement the tasks following GA, Annex 1, and will ensure that by doing so it will not knowingly infringe the third party property rights.

Management of knowledge, intellectual property, and innovation will be undertaken in Task 8.3-Exploitation strategy and IPR management (WP8 – Communication, dissemination, exploitation strategy, and action plan).

8 Ethical aspects

As stated in the GA [2](see Article 14 and Annex 5), project activities must be carried out following the highest ethical standards and applicable EU, international, and national rules on ethical aspects. TechUPGRADE partners will pay special attention to:

- Principle of proportionality
- Personal data protection rights,
- Privacy right,
- Right to the integrity of the person,
- Right to equality and non-discrimination,
- Right to a healthy environment,
- Right to a high level of health protection.

Next to this, project partners will adhere to principles of research integrity as indicated in the European Code of Conduct for Research Integrity [8], which means complying with the rules of:

- Reliability in research quality assurance is reflected in the design, methodology, analysis, and use of resources.
- Integrity in developing, undertaking, reviewing, reporting, and communicating scientific research in a transparent, honest, complete, and impartial manner.
- Respect for colleagues, research participants, society, ecosystems, cultural heritage, and the environment.
    - Responsibility for research from idea to publication, for its management and organization, for training, supervision, and mentoring, and its wider impact.

This project has received funding from European Union's Horizon Europe's Research and Innovation Program under grant agreement No. 101103966. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

CONTACT US

techupgrade.eu

FOLLOW US

Funded by
the European Union