D9.1 – Data Management Plan

Johannes Losacker & Thomas Echterhof | RWTH Aachen University

June 2023

HYINHEAT

Credits

Disclaimer

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101091456.

Technical references

Grant Agreement N°	101091456	Acronym	HyInHeat
Full Title	D9.1 Data Management Plan		
Work Package (WP)	WP9		
Authors	Johannes Losacker & Thomas Echterhof		
Leading organisation	RWTH Aachen University		
Date of publication	30/06/2023		
Contributors	All partners		
Document Type	R – Document, report		
Document Title	Data Management Plan		
Dissemination Level (delete unnecessary options)		public	

Document history

Version	Date	Partner	Author
1	30/06/2023	RWTH	JL / TE
2			
3			
4			
5			
Final			

Contents

Technical references 3
Document history 3
Contents 4
List of abbreviations 5
Executive Summary 6
HyInHeat Data Management Plan 7
Introduction 7
About the project 7
Data summary 8
FAIR data 11
Making data findable, including provisions for metadata 11
Making data accessible 12
Making data interoperable 15
Increase data re-use 16
Other research outputs 18
Allocation of resources 18
Data security 19
Ethics 19
Other issues 20
Get in touch 21

List of abbreviations

Abbreviations	Explanation
BSI	Bundesamt für Sicherheit in der Informationstechnik
CC	Creative Commons
CFD	Computational Fluid Dynamics
DOI	Digital Object Identifier
ISO	International Organization for Standardization
LCA	Life Cycle Assessment
LES	Large Eddy Simulation
MFA	Material Flow Analysis
PM	Person Month
RANS	Reynolds Averaged Navier Stokes
SSL	Secure Sockets Layer
WSGGM	Weighted Sum of Gray Gases Model

Executive Summary

This deliverable deals with the data management life cycle for the data to be collected, processed and/or generated by the HyInHeat project. The document includes information on the handling of research data during & after the end of the project, what data will be collected, processed and/or generated, which methodology & standards will be applied, whether data will be shared/made open access and how data will be curated & preserved. The key point is to collect and provide information on the data created and used in the project as well as how FAIR principles will be implemented in data management. This Data Management Plan (DMP) and its structure is based on the Horizon Europe DMP template provided by the European Commission. The Horizon Europe Model Grant Agreement requires that a data management plan is established and regularly updated. For this reason, the deliverable will be regularly updated (by D9.2 "Updated Data Management Plan" and D9.4 "Final Data Management Plan") on the basis of the project's evolution.

The document is structured in three main sections:

Section 1 - HyInHeat project presentation: Introducing the purpose of this deliverable and presenting the main objectives of the project.

Section 2 - HyInHeat data summary: provides and inventory of data handled (collected, created, re-used, …) by the HyInHeat project giving information on identified datasets, data formats, purpose of data generation or (re-)use, data size and origin/provenance of data handled.

Section 3 - HyInHeat FAIR data: provides information how the FAIR principles regarding findability, accessibility, interoperability and re-use of data are implemented in the HyInHeat project.

HyInHeat Data Management Plan

Introduction

The Horizon Europe Model Grant Agreement requires that a data management plan ('DMP') is established and regularly updated. This DMP describes the data management life cycle for the data to be collected, processed and/or generated by the HyInHeat project. The DMP includes information on the handling of research data during & after the end of the project, what data will be collected, processed and/or generated, which methodology & standards will be applied, whether data will be shared/made open access and how data will be curated & preserved. This document and its structure are based on the Horizon Europe DMP template provided by the European Commission.

The project's contents play an important role for the data to be collected, processed and/or generated by the HyInHeat project. Therefore, an introduction to the project is necessary to understand what data will be collected, processed and/or generated.

About the project

HyInHeat aims to integrate hydrogen as fuel for high temperature heating processes in the energy intensive industries. While some of the equipment is already presented as hydrogen-ready, the integration of hydrogen combustion in heating processes still needs adoption and redesign of infrastructure, equipment and the process itself.

More specifically, HyInHeat has the following objectives:

- Implement 8 demonstrators to redesign combustion systems for H_2 firing and to optimize burner technology to minimize NO_x emissions. This includes a greenfield study for steel reheating to provide a new furnace layout with an adjusted conductive zone and off-gas system.
- Modify and optimize four burners for cross-sectorial application for hydrogen operation, and develop measurement equipment for fuel supply and combustion control. H_2 compatible fuel supply lines will also be implemented, and the refractory degradation and consumption will be monitored and readjusted to maintain TRL 7 standard.
- Seven out of eight demonstrators in this project will use pure oxygen instead of air, with the potential for energy savings up to 40% through optimized combustion control systems and the use of pure oxygen. The use of pure oxygen is not yet established as the state-of-the-art in five of the seven processes being demonstrated.
- HyInHeat will use two measurement technologies for fuel quality determination and implement them into a lab-scale test rig and two demonstrators to address fuel gas characteristics and flows. Combustion control and NO_x emission measurement technologies for H_2/air and H_2/O_2 fired processes will be developed, and data analysis and data-based digital process control algorithms will be performed to promote digitalization of production processes, and these solutions will be implemented in the demonstrators to provide interdisciplinary guidelines.
- The project defines KPIs and scenario parameters to compare the developed heating technologies with fossil fuel alternatives. A baseline is set up to quantify the project's impact on KPIs such as energy consumption, GHG emissions, NO_x emissions, and metal loss. The economic viability of the H_2 heating technologies is evaluated and individual business cases are developed.

With these activities, HyInHeat contributes to the objectives of decreasing CO_2 emission of the processes while increasing energy efficiency in a cost competitive way keeping NO_x emission levels and resource efficiency at least at the same level.

Data summary

Will you re-use any existing data and what will you re-use them for?

Dataset Numerical Simulations

This dataset is created: Development of new numerical models for combustion processes and thermal radiation. Support of equipment development and furnace retrofitting. Evaluation of greenfield and retrofitting solutions.

Dataset Experimental Measurements

This dataset is created: Support of equipment development and furnace retrofitting. Comparison with and validation of numerical models. Baseline definition for new measurement technologies and NO_x emission limits Investigations on product quality, yield and refractories. Design of safe and efficient H_2 and O_2 infrastructure.

Dataset MFA/LCA

This dataset is created: Calculating the CO_2 and NO_x footprint of every industrial trial as well as characterizing the material flows from aluminium and steel.

This data is reused: Data from previous publications for characterizing the material flows from aluminium and steel

Dataset Source Code

This dataset includes both (combination of re-used and newly generated data): Measurement and control algorithm implementation. Collection of supporting tools in various fields, that cannot be foreseen yet.

Dataset Stakeholder Consultation

This dataset is created: Stakeholder's responses to questions on regulatory and market framework conditions

Dataset Communication, Dissemination & Exploitation Activities

This dataset is created: List of stakeholders deriving from D8.2, list of stakeholders registered to receive the project newsletters, contact forms submitted by website users, list of participants to project events, website cookies

State the reasons if re-use of any existing data has been considered but discarded.

Dataset Numerical Simulations

Numerical simulations make use of newly developed models and case specific geometries and hence, need to be created within the project.

Dataset Experimental Measurements

Measurements of the project specific demonstrators are not available for operation at oxygen enriched and hydrogen enriched combustion, as well as combustion of pure hydrogen with oxyfuel. Hence, the first trails and measurements under these conditions are conducted within the project.

Dataset Source Code

Existing source codes will be adjusted as well as being created from scratch.

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

Dataset Numerical Simulations

The data are - Kinetic mechanism for combustion of hydrogen/oxyfuel - Large Eddy Simulations (LES) of lab- and industrial burners - Reynolds Averaged Navier Stokes (RANS) simulations of lab- pilot- and industrial size furnaces - Weighted Sum of Gray Gases Model (WSGGM) for radiative heat transfer.

They are provided in the following formats: Table: csv, txt, xlsx; Graphics: jpeg, pdf, png, bmp; Video: mpeg, avi; Text: xml, pdf, docx, html, txt, xlsx; Database: csv, h5, xml, json; Computer Aided Design: stp.

Dataset Experimental Measurements

The data are Operational data (Temperature, pressure, density, species concentration) measurements of lab- pilot- and industrial furnaces.

They are provided in the following formats: Table: csv, txt, xlsx; Graphics: jpeg, pdf, png, bmp; Video: mpeg, avi; Text: xml, pdf, docx, html, txt, xlsx; Database: csv, h5, xml, json; Computer Aided Design: stp; Statistical analysis: dta, por, sas, sav.

Dataset MFA/LCA

The data are - Quantitative survey about emissions measurements, material input and output.

They are provided in the following formats (e.g.): Table: csv, txt, xlsx; Graphic: jpeg, pdf, png, bmp; Text: xml, pdf, docx, html, txt, xlsx; Database: csv, h5, xml, json, ecoSpold; Statistical analysis: dta, por, sas, sav.

Dataset Source Code

The data are code and software files.

They are provided in the following formats: Table: csv, txt, xlsx; Graphic: jpeg, pdf, png, bmp; Text: xml, pdf, docx, html, txt, xlsx; Database: csv, h5, xml, json.

Stakeholder Consultation

The data are responses expressing stakeholder views on policy gaps and obstacles associated with the use of H_2 in industrial heating processes identified in a stakeholder consultation.

They are provided in the following formats: table: csv, txt, xlsx; audio: wave, aiff, mp3, mxf, flac; text: xml, pdf, docx, html, txt, xlsx.

Communication, Dissemination & Exploitation Activities

The data are Website cookies Newsletter subscriptions Contact forms.

They are provided in the following formats: table: csv, txt, xlsx; text: xml, pdf, docx, html, txt, xlsx; database: csv, h5, xml, json.

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

Dataset Numerical Simulations

Development of new numerical models for combustion processes and thermal radiation. Support of equipment development and furnace retrofitting. Evaluation of greenfield and retrofitting solutions.

Dataset Experimental Measurements

Support of equipment development and furnace retrofitting. Comparison with and validation of numerical models. Baseline definition for new measurement technologies and NOx emission limits. Investigations on product quality, yield and refractories. Design of safe and efficient H2 and O2 infrastructure.

Dataset MFA/LCA

Calculating the CO2 and NOx footprint of every industrial trial as well as characterizing the material flows from aluminium and steel.

Dataset Source Code

Measurement and control algorithm implementation. Collection of supporting tools in various fields, that cannot be foreseen yet.

Dataset Stakeholder Consultation

Derivation of policy recommendations

Dataset Communication, Dissemination & Exploitation Activities

Public communication, exploitation of project results and website operation

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

Dataset Numerical Simulations

between 1 TB and 100 TB

Dataset Experimental Measurements

between 1 TB and 100 TB

Dataset MFA/LCA
between 1 GB and 1 TB

Dataset Source Code
between 1 GB and 1 TB

Dataset Stakeholder Consultation
between 1 KB and 1 GB

Dataset Communication, Dissemination & Exploitation Activities
between 1 KB and 1 GB

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

Dataset Numerical Simulations

- Furnace geometries for CFD simulations are either provided by demonstrator furnace operators or generated from measurements and construction data.
- Burner geometries are provided by operators or burner manufacturers.
- CFD simulations will be set up by the partners BSC and RWTH.

Dataset Experimental Measurements

- Industrial measurements are performed at the respective demonstrator's facilities.
- Lab measurements are performed at the research centres and universities involved in the project.

Dataset MFA/LCA

- Datasheet from industrial partners from the trials.
- EcoInvent database.
- Literature

Dataset Source Code

- Measurement and control algorithms will be provided/developed by project partner SICK.
- Further source code can be contributed potentially by any partner.

Dataset Stakeholder Consultation

- Key stakeholders fill out a questionnaire

Dataset Communication, Dissemination & Exploitation Activities

- All public communication activities, e.g. interactions on the project website, newsletter subscriptions, contact form use

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

Dataset Numerical Simulations

- Other scientists (from the own or another field): Scientists in the field of hydrogen combustion.
- Stakeholders from industry / economy: Burner manufacturers, furnace operators.

Dataset Experimental Measurements

- Other scientists (from the own or another field): Scientists in the field of hydrogen and oxyfuel combustion.
- Stakeholders from industry / economy: Burner manufacturers, furnace operators.
- Standardisation bodies: Bodies in charge of emission limit definition.
- Politics / Decision makers: Environmental policy actors.

Dataset MFA/LCA

- Other scientists (from the own or another field): Steel and Secondary aluminium sector.
- Stakeholders from industry / economy: Steel and Secondary aluminium sector.
- Politics / Decision makers: Environmental policy actors.

Dataset Source Code

- Other scientists (from the own or another field): Scientists in the field of hydrogen and oxyfuel combustion.
- Stakeholders from industry / economy: Burner manufacturers, furnace operators.

Dataset Stakeholder Consultation

- Stakeholders Industry and Developers: Aluminium and Steel industry, Furnace and Equipment Developers, Refractory Developers, Combustion & Measurement Equipment Developers
- Stakeholders Suppliers and Researchers: Hydrogen Suppliers, Oxygen Suppliers, Research & Solution Developers, Steel and Aluminium Users, Policy Actors, Other Process Industries
- Stakeholders Wider public: Wider Public, Energy Suppliers & Other Raw Materials Providers

FAIR data

Making data findable, including provisions for metadata

Will the data be identified by a persistent identifier?

Dataset Numerical Simulations

Yes, Handle / DOI

Dataset Experimental Measurements

Yes, Handle / DOI

Dataset MFA/LCA

Yes, Handle / DOI

Except for confidential data

Dataset Source Code

No

Dataset Stakeholder Consultation

No

Dataset Communication, Dissemination & Exploitation Activities

No

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.

The data will be described using the following metadata:

Dataset Numerical Simulations

- Automatically created: Timestamp, Storage.
- Automatically created, manually corrected: Title, Author.
- Manually created: to be clarified with repository.

Dataset Experimental Measurements

- Automatically created: Timestamp, Storage.
- Automatically created, manually corrected: Title, Author.
- Manually created: to be clarified with repository.

Dataset MFA/LCA

- Automatically created: None.
- Automatically created, manually corrected: None.
- Manually created: to be clarified with repository.

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

Dataset Numerical Simulations

Yes

Dataset Experimental Measurement

Yes

Dataset MFA/LCA
Yes

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

Dataset Numerical Simulation
to be clarified with repository

Dataset Experimental Measurement
to be clarified with repository

Dataset MFA/LCA
to be clarified with repository

Making data accessible

Repository

Will the data be deposited in a trusted repository?

Dataset Numerical Simulation
The dataset is stored at: own institution (e.g., RWTH Coscine); To be decided.

RWTH Coscine repository is trusted because:

- The repository meets the basic expectations of a research data repository: easy and free access while observing legal and ethical barriers, detailed metadata, protection against unauthorized access.
- Machine-actionable and standardised metadata are used.
- The repository can provide persistent identifiers: DOI.
- The repository offers expert curation for the long-term archiving.
- The repository is managed by a trustworthy institution: RWTH Aachen University.
- The repository's policy is available online: https://coscine.rwth-aachen.de/login/?redirect=%2F

Dataset Experimental Measurement
The dataset is stored at: own institution (e.g., RWTH Coscine); To be decided.

RWTH Coscine repository is trusted because:

- The repository meets the basic expectations of a research data repository: easy and free access while observing legal and ethical barriers, detailed metadata, protection against unauthorized access.
- Machine-actionable and standardised metadata are used.
- The repository can provide persistent identifiers: DOI.
- The repository offers expert curation for the long-term archiving.
- The repository is managed by a trustworthy institution: RWTH Aachen University.
- The repository's policy is available online: https://coscine.rwth-aachen.de/login/?redirect=%2F

Dataset MFA/LCA
The dataset is stored at: own institution (e.g. RWTH Coscine); To be decided.

RWTH Coscine repository is trusted because:

- The repository meets the basic expectations of a research data repository: easy and free access while observing legal and ethical barriers, detailed metadata, protection against unauthorized access.
- Machine-actionable and standardised metadata are used.
- The repository can provide persistent identifiers: DOI.
- The repository offers expert curation for the long-term archiving.
- The repository is managed by a trustworthy institution: RWTH Aachen University.
- The repository's policy is available online: https://coscine.rwth-aachen.de/login/?redirect=%2F

Dataset Source Code
The dataset is stored at: own institution (e.g., GitLab RWTH); To be decided.

Dataset Stakeholder Consultation
The dataset is stored at: own institution; To be decided.

Dataset Communication, Dissemination & Exploitation Activities
The dataset is stored at: own institution.

Have you explored appropriate arrangements with the identified repository where your data will be deposited?
Dataset Numerical Simulations
Yes, because of the big size of the data.

Datasets Experimental Measurements, MFA/LCA, Source Code
No, the data will be uploaded via a standard procedure and requires no special arrangements.

Does the repository ensure that the data are assigned an identifier?
Dataset Numerical Simulations
Yes

Dataset Experimental Measurements
Yes

Dataset MFA/LCA
Yes

Dataset Source Code
No

Dataset Stakeholder Consultation
No

Dataset Communication, Dissemination & Exploitation Activities
No

Will the repository resolve the identifier to a digital object?
Dataset Numerical Simulations
Yes, the repository will resolve the identifier to a digital object.

Dataset Experimental Measurements
Yes, the repository will resolve the identifier to a digital object.

Dataset MFA/LCA
Yes, the repository will resolve the identifier to a digital object.

Data:

Will all data be made openly available?
Datasets Experimental Measurements, MFA/LCA, Source Code
Yes, internally with everyone, as long as they don't publish the data or share it externally.

Datasets Stakeholder Consultation, Communication Activities
No

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.

Dataset Numerical Simulations
Legal Restrictions: Parts of the data are intellectual property of a partner or third party and the project does not own the right to share them.

Dataset Experimental Measurements
Legal Restrictions: Permission is required to access the data or the object being researched and the grantors of the permission do not consent to disclosure of the data.

Dataset MFA/LCA
Legal Restrictions: Permission is required to access the data or the object being researched and the grantors of the permission do not consent to disclosure of the data.

Dataset Source Code
Legal Restrictions: Parts of the data are intellectual property of a partner or third party and the project does not own the right to share them.

Dataset Stakeholder Consultation
Name or any other personal identifying information will not appear in any publications resulting from the responses to the consultation; neither will there be anything to identify the respondents' place of work or the projects they're involved in.

Dataset Communication, Dissemination & Exploitation Activities
Legal restrictions: Permission is required to access the data and the handling of them will be in compliance with GDPR requirements. For more information, please visit https://hyinheat.eu/privacy-statement/

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.

Dataset Numerical Simulations
An embargo is applied for

- Publication of the results in a peer-reviewed journal.
- Achievement of intellectual property rights such as: Patents.

Dataset Experimental Measurements
An embargo is applied for

- Publication of the results in a peer-reviewed journal.
- Achievement of intellectual property rights such as: Patents.

Dataset MFA/LCA
An embargo is applied for

- Publication of the results in a peer-reviewed journal.

Dataset Source Code, Stakeholder Consultation, Communication Activities
No embargo is applied.

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

The partners are encouraged to choose one of the following Creative Commons licenses:

- Creative Commons Attribution (CC-BY)
  - Permission for the work to be distributed, adapted, and build upon, even commercially while crediting the author for the original creation.
- Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)
  - Permission for the work only to by accessed and distributed while crediting the author for the original creation. Altering and commercial use of the article is not permitted.

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

Dataset Numerical Simulations
Furnace and burner geometries will be hold back partly, because open publication could harm the respective participants market position.

Dataset Experimental Measurements
The dataset might contain sensitive data about involved equipment that is not meant to be shared publicly due to marked competitiveness considerations.

Dataset Source Code
Public sharing of measurement and control algorithms might harm the partners marked position.

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

If necessary, an authentication system or a data on demand function will be provided.

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

This consortium has not established a Data Access Committee. The appointed data responsible / corresponding author will decide alone about granting access to the data.

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.

Dataset Numerical Simulations, Experimental Measurements, MFA/LCA, Source Code
Yes

Dataset Stakeholder Consultation
The data will not be shared publicly

Dataset Communication, Dissemination & Exploitation Activities
The data will not be shared publicly

Will metadata contain information to enable the user to access the data?

The need for special software will be stated in the metadata.

How long will the data remain available and findable?

10 years

Will metadata be guaranteed to remain available after data are no longer available?

No.

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)?

The need for special software will be stated in the metadata.

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?

Dataset Numerical Simulations
Discipline-specific standards, classifications etc. are used.

Dataset Experimental Measurements
Discipline-specific standards, classifications etc. are used.

Dataset MFA/LCA
Discipline-specific standards, classifications etc. are used: e.g. ISO 14044 and 14040.

Dataset Source Code
Discipline-specific standards, classifications etc. are used.

Will you follow community-endorsed interoperability best practices? Which ones?

Dataset MFA/LCA
Standardized formats are used for the data set: e.g. ISO 14044 and 14040.

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?

No mappings are necessary, as the datasets are described using standard terminologies / There is no need for a mapping, since the used terminology is chosen to be compatible with the existing literature.

Will your data include qualified references[1] to other data (*e.g.,* other data from your project, or datasets from previous research)?

Dataset Numerical Simulations

Yes, the data include qualified references to other datasets from the same project: Kinetic mechanism, Measurement data.

Dataset MFA/LCA

Yes, the data include qualified references to other datasets from the same project: Data collection from demonstrators.

Dataset Source Code

Yes, the data include qualified references to other datasets from the same project: Data collection from demonstrators.

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.)?

Dataset Numerical Simulations

We will provide documentation respective repository, published journal articles, project reports, in the form of:

- Information on methodology used for data generation (creation or re-use) and transformation: Information on models, meshing, boundary conditions and demonstrator cases are provided.
- Information on methodology used for data cleaning and analysis: Post processing follows established procedures.
- Description of variables used and respective units of measurement: Variables and units are documented in respective reports.
- Version information.

Dataset Experimental Measurements

We will provide documentation respective repository, published journal articles, project reports, in the form of:

- Information on methodology used for data generation (creation or re-use) and transformation: Measurement equipment and measurement conditions are documented.
- Information on methodology used for data cleaning and analysis: Follows common procedures and is documented.
- Description of variables used and respective units of measurement: Variables and units are documented in respective reports.
- Version information.

Dataset MFA/LCA

We will provide documentation respective repository, published journal articles, project reports, in the form of:

[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/)

- Information on methodology used for data generation (creation or re-use) and transformation.
- Information on methodology used for data cleaning and analysis.
- Description of variables used and respective units of measurement.

Dataset Source Code

We will provide documentation respective repository, published journal articles, project reports, in the form of:

- Description of variables used and respective units of measurement.
- Version information.

Will your data be made freely available in the public domain to permit the widest re-use possible?

For datasets not mentioned below: Yes, internally with everyone, as long as they don't publish the data or share it externally.

Dataset Communication, Dissemination & Exploitation Activities

No

Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

The partners are encouraged to choose one of the following Creative Commons licenses:

- Creative Commons Attribution (CC-BY)
  - Permission for the work to be distributed, adapted, and build upon, even commercially while crediting the author for the original creation.
- Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)
  - Permission for the work only to by accessed and distributed while crediting the author for the original creation. Altering and commercial use of the article is not permitted.

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

The data will be available for re-use, as far as reported above.

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

Yes

Describe all relevant data quality assurance processes

Dataset Numerical Simulations

- Completeness check.
- Data reconciliation.
- Repeated or comparative measurements.
- Use of controlled vocabularies and standard terminology.
- Data validation: Comparison to demonstrator measurements.
- Provision of test results along with the data.
- Peer review of textual publications based on the data.

Dataset Experimental Measurements

- Completeness check.
- Data reconciliation.
- Repeated or comparative measurements.
- Adherence to standard procedures for data recording.
- Use of controlled vocabularies and standard terminology.
- Peer review of textual publications based on the data.

Dataset MFA/LCA

- Completeness check.
- Use of controlled vocabularies and standard terminology.

- Data validation.
- Peer review of textual publications based on the data.

Dataset Source Code

- Completeness check.
- Use of controlled vocabularies and standard terminology.

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.

We will address these aspects in the following section.

Other research outputs

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.).

The project will produce research output also of the following types:

- Workflows
- Models
- Samples
- Other: Policy recommendations regarding emission limit definition

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.

Listed other research outputs cannot be uniquely assigned to one of the dataset-categories, but rather are more general outcomes of the project. While the underlying research is covered by the dataset categories, outcomes as workflows, models and policy recommendations will be derived and documented in respective project reports and where appropriate, will be covered by textual publications.

Allocation of 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.)?

The following data is a first estimation that will be updated with the following versions of this document.

For metadata: 12 PM ; 0 Euro

For PIDs: 3 PM ; 0 Euro

For data preservation: 5 PM ; 0 Euro

For management of other research outputs: 10 PM ; 0 Euro

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).

Data management costs will be covered by the Horizon Europe grant.

Who will be responsible for data management in your project?

RWTH: Johannes Losacker

TEC: Aitziber Adrados

OULU: Henri Pauna

AGA: Esin Iplik

EGEN: Thomas Maidonis

Further responsible persons will be added in updated versions of this DMP.

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).

Duration of preservation: 10 years.

Costs: see above.

Potential value: see above.

Criteria for preservation:

- Basis for textual publication
- potential re-use
- legal or contractual conditions
- quality considerations
- reproducibility

Motivation for long-term preservation

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)?

RWTH Coscine

- Protection against unauthorized access: all systems are operated according to BSI IT-Grundschutz
- Data robustness: encryption: SSL-encryption
- Data recovery: backups: Metadata backup
- Sensitive data: anonymisation or pseudonymisation

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

Yes, the chosen repositories will be trustworthy. See above for further details.

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).

Personal data

Dataset Stakeholder Consultation

The dataset contains personal data. Some data contain information allowing person identification. The personal data will be anonymised/pseudonymised: after the data analysis / before publication.

Dataset Communication, Dissemination & Exploitation Activities

The dataset contains personal data. Some data contain information allowing person identification. The personal data will be anonymised/pseudonymised: No.

Other sensitive data

Dataset Numerical Simulations

The geometries contained in the dataset are in part subject to business secrets.

Dataset Experimental Measurements

Burner and furnace measurements might reveal operation strategies, which fall into the category of business secrets.

Dataset MFA/LCA
Technical drawings and specific data from facilities (production data, photos, machines...) are considered as business secret.

Dataset Stakeholder Consultation
The stakeholder consultation can involve information that is considered confidential for some of the partners.

Intellectual property rights

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

Dataset Stakeholder Consultation
Yes, only for analysis / use of the data within the project.

Dataset Communication, Dissemination & Exploitation Activities
Yes, only for analysis / use of the data within the project.

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)?

Regulations of the involved partner institutions
OULU: optional - CSC's IDA/ALLAS service (national Finnish data management tool)

Project-internal regulations
Metadata is collected using a standardized data definition document. Data definition includes

- An overview: Related project work packages and tasks, storage location and data access rights
- Work description: Author, date, location, purpose of the work/data, experimental setup, equipment, work steps, supplementary material
- Data description: Data file types and formats, potential usage, data analysis or modification, associated software, variable description

Get in touch

Website

http://hyinheat.eu/

Email address

info@hyinheat.eu