Project Title: **Industrial Thermal Energy Recovery conversion and Management**
Project Acronym: **I-ThERM**
Grant Agreement No: **680599**

D11.3 Data Management and Open Access Strategy
February 2018

Deliverable leader	Thomas Parissis (CETRI)
Contributors	
Hussam Jouhara (BRUNEL), Savvas Tassou (BRUNEL), Yunting Ge (BRUNEL), Lesley Hanna (BRUNEL), Andrew Holgate (Econotherm), Mark Boocock (Econotherm), Arthur Leroux (Enogia), Maxence de Miol (Enogia), Benoit Pailette (Enogia), Hans Schweiger (Energyxperts), Claudia Vannoni (Energyxperts), Steven Woolass (TATA), Stuart Bennett (TATA), Fran Godley (TATA), Francisco-Javier Lago (ArcelorMittal), Ramon Laso Ayuso (ArcelorMittal), Rocio Llera Traviesa (ArcelorMittal), Franco Antonio Cavadini (Synesis), Giuseppe Montalbano (Synesis), Paul Christodoulides (CUT), Soteris Kalogirou (CUT), George Florides (CUT), Gregoris Panayiotou (CUT), Javier Gomez (Avanzare), Javier Perez (Avanzare), Marta Perez (Avanzare), Julio Gomez (Avanzare), Jeremy Miller (SpiraxSarco), David Oliver (SpiraxSarco), Obadah Zaher (SpiraxSarco), Jesus Pavia (Arluy), Vassilis Stathopoulos (TEISTE)	

Dissemination Level		
PU	Public, fully open, e.g. web	X
CO	Confidential, restricted under conditions set out in Grant Agreement	

	Version Control	
1	*Version 1*	*28/12/2015*
2	*Version 2*	*22/01/2016*
3	*Version 3*	*17/02/2016*
4	*Version 4*	*23/05/2016*
5	*Version 5*	*11/07/2016*
6	*Version 6*	*28/02/2018*

DISCLAIMER

This document contains the description of the I-ThERM project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules. Therefore, prior to using its content please contact the consortium coordinator for approval. E-mail: Ithermproject@gmail.com

Should you feel that this document harms in any way the IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken all available measures in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This document has been produced with the assistance of the European Union. The content of this document is the sole responsibility of the I-ThERM consortium and can in no way be taken to reflect the views of the European Union.

Contents

Executive Summary

The purpose of this document is to set the Data Management and Open Access strategy for the I-ThERM project. It contains guidelines that will be used for the development of a Data Management Plan (DMP) which will include an analysis of the main elements of the data management policy that will be used by the I-ThERM consortium with regards to all the data that will be generated by the project. Moreover the DMP will cover the following aspects:

- Description of the data to be collected / created
- Standards / methodologies for data collection and management
- Ethics and Intellectual Property concerns or restrictions
- Plans for data sharing and access
- Strategy for long-term preservation

The DMP will not be a fixed document, but it will evolve and will will gain more precision and substance during project implementation. New versions of the DMP will be created whenever important changes to the project occur due to inclusion of new data sets, changes in consortium policies or external factors. The first version of the DMP is expected to be delivered within the first months of the project when the first data sets are identified. More detailed versions of the DMP will be delivered at later stages of the project. The DMP would need to be updated at least by the mid-term and final review to fine-tune it to the data generated and the uses identified by the consortium since not all data or potential uses are clear from the start.

1   Introduction

The overall aim of I-ThERM project is to develop and demonstrate technologies and processes for efficient and cost effective heat recovery from industrial facilities in the temperature range 70 $^{\circ}$C to 1000 $^{\circ}$C and the optimum integration of these technologies with existing energy systems or for over the fence export of recovered heat and generated electricity if appropriate.

To achieve this challenging aim, and ensure wide application of the technologies and approaches developed, the project brings together a very strong consortium comprising of RTD providers, technology providers and more importantly large and SME users who will provide demonstration sites for the technologies. The project will focus on two phase innovative heat transfer technologies (heat pipes-HP) for the recovery of heat from medium and low temperature sources and the use of this heat:

 a) within the same facility or export over the fence;
b) for generation of electrical power; or
c) a combination of (a) and (b) depending on the needs.

For power generation the project will develop and demonstrate at industrial sites the Trilateral Flash System (TFC) for low temperature waste heat sources, 70 $^{\circ}$C to 200 $^{\circ}$C and the Supercritical Carbon Dioxide System (sCO2) for temperatures above 200 $^{\circ}$C. It is projected that these technologies used alone or in combination with the HP technologies will lead to energy and GHG emission savings well in excess of 15% and attractive economic performance with payback periods of less than 3,0 years.

The long term objectives of this three year project are the following:

1) Identify and quantify streams of waste heat from industrial processes in the EU 28 and potential for energy recovery
2) Use the 'EINSTEIN' toolkit to carry out energy audits, analyse the technical potential and economic viability of heat recovery.
3) Develop heat recovery technologies and equipment in packaged or easily customisable plug and play forms that can readily be selected and applied in industry.

4) Develop an intelligent system for monitoring and on-line integration and control of the operation of these technologies to maximise heat recovery and minimise operating costs and emissions.

5) Investigate and evaluate organisational, technoeconomic and socioeconomic barriers to the wide adoption of advanced heat recovery technologies and ways of overcoming these barriers.

6) Implement, monitor and evaluate the performance of heat recovery applications, evaluate their impact on overall energy consumption and CO_2 emissions and disseminate the outputs widely to industry, other key stakeholders and policy makers.

The purpose of this document is to set the Data Management and Open Access strategy for the I-ThERM project. It will include guidelines that will be used for the development of a Data Management Plan (DMP) which will include an analysis of the main elements of the data management policy that will be used by the I-ThERM consortium with regards to all the data that will be generated by the project. Moreover the DMP will cover the following aspects:

- Description of the data to be collected / created
- Standards / methodologies for data collection and management
- Ethics and Intellectual Property concerns or restrictions
- Plans for data sharing and access
- Strategy for long-term preservation

The DMP will not be a fixed document, but it will evolve and will be further enriched during project implementation. The DMP will be updated at least by the mid-term and final review to fine-tune it to the data generated and the uses identified by the consortium since not all data or potential uses are clear from the start. New versions of the DMP will be created whenever important changes to the project occur due to inclusion of new data sets, changes in consortium policies or external factors.

2   The Data Management Plan

2.1   Purpose of the Data Management Plan.

The Data Management Plan (DMP) purposes are:
- to support the data management life cycle for all data that will be collected, processed or generated by the Project,

- to provide an analysis of the main elements of the data management policy that will be used by the applicants with regard to all the datasets that will be generated by the Project,
- to provide detail and guarantee about the preservation of the data collected during the Project, as well as any results derived from the associated research,
- to provide detail on how we plan to address the Ethical issues (if any) related to data that will be collected during the Project timeframe,
- to create a document which explains the management of data collected during the Project.

The DMP is not a fixed document, but it will evolve during the Project.

2.2   Intended audience

The DMP will be oriented to:

- the project's participant organizations;
- the local Ethics Committee;
- the partners's personnel and all stakeholder involved by the Project;
- the European Commission.

2.3   Overview of the DMP

The DMP will contain details on:

- a brief description of data types which will be collected during the I-ThERM project, explaining the procedures used to collect or create them;
- copyright and IPR issues;
- ethical issues related to data storage, persons authorized to see/use them and how long they are kept; managing ethical concerns that include the anonymization of data; procedures used to obtain the consent requested to allow data sharing and reuse.

2.4   Terminology

2.4.1   Abbreviations and acronyms

- EC : European Commission.
- EU: European Union
- OMB: Office of Management and Budget
- DMP: Data Management Plan

- WP: Work Package

2.4.2   Definitions

Words beginning with a capital letter shall have the meaning defined either herein or in the Rules or in the Grant Agreement related to the Project;

2.4.3   Additional Definitions

- **Project**: Project refers to the I-ThERM project funded from the European Union's Horizon 2020 research and innovation programme under Grant Agreement 680599.

- **Metadata**: Metadata is data that describes other data. Meta is a prefix that in most information technology usages means "an underlying definition or description." Metadata summarizes basic information about data, which can make finding and working with particular instances of data easier.

3   Data Sets

3.1   Defining research data

One definition of research data is: "the recorded factual material commonly accepted in the scientific community as necessary to validate research findings." (Office of Management and Budget (OMB), Uniform Administrative Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations, **CIRCULAR A-110 REVISED 11/19/93 As Further Amended 9/30/99,** OMB Circular 110). Research data covers a broad range of types of information, and digital data can be structured and stored in a variety of file formats. Note that properly managing data (and records) does not necessarily equate to sharing or publishing that data.

Some examples of research data include:
- Documents (text, Word), spreadsheets
- Laboratory notebooks, field notebooks, diaries
- Questionnaires, transcripts, codebooks
- Audiotapes, videotapes
- Photographs, films
- Protein or genetic sequences
- Spectra

- Test responses

- Slides, artifacts, specimens, samples

- Collection of digital objects acquired and generated during the process of research

- Database contents (video, audio, text, images)

- Models, algorithms, scripts

- Contents of an application (input, output, logfiles for analysis software, simulation software, schemas)

- Methodologies and workflows

- Standard operating procedures and protocols

In addition to the other records to manage, some kinds of data may not be sharable due to the nature of the records themselves, or to ethical and privacy concerns. As defined by the OMB (Office of Management and Budget (OMB), Uniform Administrative Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations, **CIRCULAR A-110 REVISED 11/19/93 As Further Amended 9/30/99,** OMB Circular 110), this refers to:

- preliminary analyses,

- drafts of scientific papers,

- plans for future research,

- peer reviews, or

- communications with colleagues

Research data also do not include:

- Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and

- Personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.

The following research records may also be important to manage during and beyond the life of a project.

- Correspondence (electronic mail and paper-based correspondence)

- Project files

- Grant applications

- Ethics applications

- Technical reports
- Research reports
- Signed consent forms

3.2   Data sets

The specific Data Sets for the I-ThERM project need to be identified and described with the contribution of all the project partners. A short description of the data that will be generated in the research project (e.g., samples, physical collections, software, curriculum materials, and other materials to be produced during the course of the project) must be provided. Additionally an estimation of the amount of data and content of the data (if possible) must be included.

For this reason tables will be circulated among partners by the WP leader in order to collect information regarding data sets. The tables will include for each data set information regarding Data Set reference and Name, Description of Data sets, Standards, Metadata, Data Sharing information and information regarding Archiving and preservation (including Storage and BackUp)

All the partners will be asked to provide information regarding the data that will derive from the Work Packages and the Tasks they are leading. For the Data sets that will be identified, all the partners will need to provide adequate information regarding the following issues:

Are you generating the data or sourcing it from somewhere else under certain terms and conditions?
Is the data digital or non-digital, or both?
How will the data be created or collected? What instruments or tools will be used to produce the data?
What transformations will the data undergo? What software or file formats will you use as you work with the data?
Will the data be updated or become redundant as you make revisions and produce subsequent versions?
Is the data sensitive or confidential?
Is there ethics approval, or is ethics approval required?

From the information that will be gathered the roles of the partners and the use of the data will be identified. As a result for each type of Research data it will be defined who will be providing the data and who will be using/analyzing the data according to the following table.

Type of Research Data	Who is providing the data	Who is using/analysing the data

Additionally, the file formats that will be used is an important issue. The formats that will be used should be the best for long-term preservation and continued access of data. Formats most likely to be accessible in the future are:

- Non-proprietary and not tied to a specific piece of software
- Open, documented standard
- Common, used by the research community
- Standard representation (ASCII, Unicode)
- Unencrypted
- Uncompressed

3.3   Descriptive information and Metadata

In the DMP it will be defined what documentation and metadata will accompany the data. Metadata is structured information describing the characteristics of a resource; for example, the dates associated with a dataset or the title and author of a book. Metadata supports discovery, re-use and long-term preservation of resources. Metadata needs vary across scientific fields, but typically would cover the following:

- general descriptive and access metadata,

- data characteristics,

- archive terms and access policies.

A metadata record consists of a set of predefined elements that define specific attributes of a resource. Each element can have one or more values; for example, a dataset may have multiple creators. Documenting data enables other researchers to discover your data. Metadata about the nature of your files is also critical to the proper management of digital resources over time.

 All the partners will agree on specific issues regarding for example:

- The way that the data will be organised or formatted so that everyone working on it now and in the future knows the origins of the data.

- The way that the each file will be named (File Naming Conventions). The use of the following format is proposed and is approved by the coordinator for each file/document: "Date (yyyymmdd)_project_company_filename_author_version". For example, the file containing the minutes from the kick-off meeting is called: '20151009_IThERM_UBRUN_Minutes of Kick Off Meeting_SAT_Final'.

- Providing adequate metadata within the dataset (e.g. field labels or column headings) in order to be easy to interpret the data. Other examples of information that the data will need to contain include:
  - ✓ reference period,
  - ✓ Project funding information: European Union logo and information about Grant Agreement and the action/program that funds the project
  - ✓ Release policy including dissemination rules and purposes
  - ✓ Information about data collection (source, frequency and adjustments)
  - ✓ Keywords (Keywords or phrases describing the subject or content of the data)
  - ✓ Geographic coverage of the dataset (if applicable)
  - ✓ File formats
  - ✓ Comments

- Ways to identify different versions. It is proposed in each data set to include a versioning table, additionally to use the prefix ".v1" in each file/document name relevant to the versioning table. For versioning the rule that will be followed will be the use of a sequentially numbered system: v1, v2, v3, etc and "Final" for the final version. If changes need to be done in the final version then the name of the document will change including the relevant sequential version number, ensuring that the document with the "Final" prefix is indeed the final one.

At a minimum, metadata records should be kept in a fielded form, such as a spreadsheet, CSV file, or tab-delimited file. Auxiliary information necessary to interpret the metadata - such as explanations of codes, abbreviations, or algorithms used - should be included as accompanying documentation.

The Data sets identified for the I-ThERM project from each work package are included in Annex 1.

3.4   Ownership (IPR)

In the DMP issues regarding copyright and Intellectual Property Rights of the data will be included. This issues are set in the Consortium Agreement and the Grant Agreement of the I-ThERM project regarding all the results of the project. Thus the DMP will follow the Consortium Agreement and the Grant Agreement that is signed by all project partners regarding Ownership issues.

Materials generated under the I-ThERM Project will be disseminated in accordance with Consortium Agreement. Those that use the data (as opposed to any resulting manuscripts) shall cite and annotate it as follows:

The data were created by the I-ThERM project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 680599. For reuse of this data, please, contact I-ThERM Consortium: www.itherm-project.eu, ithermproject@gmail.com . Include your proposed use of the data to assist us in determining your eligibility and to help us navigate possible conflicts between research projects. We will provide you with a short data sharing agreement for you and your authorized institutional official to sign prior to your receiving the data.

This information will be described also in the metadata.

3.5   Storage and access

Issues regarding the way that the data will be stored and backed up will be also described in the DMP.

To ensure the safety of the data, the involved participants will use their available local file servers to periodically create backups of the relevant materials.

Additionally, all other relevant documentation created during the project such as deliverables will be self - archive and preserved in a Google Drive folder that has been created for the purposed of the project. Google Drive is a free file storage and synchronization service created by Google. It allows users to store files in the cloud, share files, and edit documents, spreadsheets, and presentations with collaborators.

The I-ThERM Google Drive Folder will be accessible by all of the partners of the I-ThERM consortium and it will include the following subfolders/architecture:

- Deliverables (where all the deliverables will be stored – drafts in the beginning that will be substituted by the finals when submitted)
- Financial (including documents regarding financial issues of the project)
- General (including general documents of I-ThERM project eg. Grant Agreement, Consortium Agreement etc.)
- Meeting (including folders for each meeting containing relevant documents and data)
- Periodic Report (Including data from the periodic reporting)
- WPs (containing one folder for each WP (11 in total) in which data relevant to each WP will be included)
- Other (to store and share any other document relevant to the I-ThERM project)

A third level of storage and accessibility will be the members section in the I-ThERM website (Private documents). In this section, the partners will be granted a username and password in order to access the folder of the private documents. The subfolders architecture will follow the Google Drive I-ThERM folder architecture that was described before.

All of the research data and material will be in place for at least the 2 years after the end of the project prescribed by the European Commission, as well as the foreseeable future following that according to the agreements reached by the consortium by the end of the project (if any additional is agreed).

The Coordinator of the I-ThERM project along with the Dissemination & Exploitation Manager will be in charge for data management and all the relevant issues.

4   Ethical and Legal issues

4.1   Ethical Issues

The I-ThERM partners are to comply with the ethical principles as set out in Article 34 of the Grant Agreement, which, among other, states that all activities must be carried out in compliance with:

(a) ethical principles (including the highest standards of research integrity — as set out, for instance, in the European Code of Conduct for Research Integrity — and including, in particular, avoiding fabrication, falsification, plagiarism or other research misconduct) and

(b) applicable international, EU and national law.

4.2   Confidentiality

All I-ThERM partners must keep any data, documents or other material confidential during the implementation for the project and for four years after end of the project in accordance with Article 36 of the Grant Agreement. Further detail on confidentiality can be found in Article 36 of the Grant Agreement.

5   Conclusion

The document presented the Data Management and Open Access strategy for the I-ThERM project. The Initial I-ThERM Data Management Plan will be developed after the identification of the initial Data sets by all partners. The DMP will be revised and updated during the entire duration of the project. An update of the Data Management Plan will be provided in the Interim report to the EU on Month 24 of the project (September 2017). The DMP will be updated at least by the mid-term and final review to fine-tune it to the data generated and the uses identified by the consortium since not all data or potential uses are clear from the start. New versions of the DMP will be created whenever important changes to the project occur due to inclusion of new data sets, changes in consortium policies or external factors.

Disclaimer

6   ANNEXES

6.1   ANNEX 1. I-ThERM Data Sets

No	Data set name	Standards	Contributors
1	Results of identification and quantification of primary energy consumption in the major industrial sectors, waste streams, conventional candidate technologies for energy recovery from these waste streams (advantages, disadvantages and limitations). Literature review of energy use and potential for heat recovery in the EU28. Detailed report that includes comparative tables with regard to country, type of industry, technology used, energy waste stream type and potential.	Word, Report	CUT, Tata Steel, EXP, ARLUY, Spirax
2	EINSTEIN thermo economic Data from detailed energy audits on the sites in Tata Steel and AME	Word, Excel, Designs	CUT, AME, Spirax, Etherm, UBRUN, Enogia
3	Determination of projected energy and CO2 emission savings from the application of the proposed heat recovery technologies in the EU28 - definition of target energy performance and capital and installation costs. Report on estimation of energy, environmental and economic potential for heat recovery in EU28	Word, Report	CUT, AME, Spirax, Etherm, UBRUN, Enogia
4	Report on barriers to the adoption of heat recovery technologies and recommendations on how to overcome them	Word, Report	CUT, UBRUN, Tata Steel, AME,EXP, CETRI
5	Definition of the specifications and requirements for the EINSTEIN tool	Word, Excel, Designs	EXP
6	Develop detailed design and simulation models (calculation and design algorithms, data bases) for heat recovery and waste heat to power technologies	Word, Excel, Designs	EXP
7	Continuous time monitoring & predicted data from EINSTEIN tool (heat demands and available waste heat, intermediate heat flows of utilities) – New versions of EINSTEIN softeware	Report	EXP
8	System control parameters (e.g. flow rates and in heat recovery systems, etc.).	Word, Excel, Designs	EXP
9	Detailed design of the TFC system for the identified application (design parameters) - Parametric analysis TFC system	Matlab	UBRUN, SPIRAX
10	Working fluids for TFC system	MS excel	UBRUN, SPIRAX
11	TFC System modelling (system performance & output, ambient temperature variation, shift patterns) -Performance maps of TFC twin-screw expander	Matlab	UBRUN, SPIRAX
12	Data from Building & testing TFC system (heat loads)	Word, Excel, Designs	SPIRAX
13	Data from sCO2 Overall system design and modelling (different design configurations for power generation from waste heat sources, performance and efficiencies of individual components, change in operating parameters such as pressure and temperature during heat addition and heat rejection) - Numerical simulations from thermodynamic modelling	Matlab	ENOGIA, UBRUN
14	Numerical simulations from thermodynamic modelling	Matlab	UBRUN

15	Detailed CAD model of the Compressor-Generator-Turbine "CGT" unit as well as electrical schematics and all other needed documents for "CGT" manufacturing, CFD model data (rotating speed and shaft power)	PDF report	ENOGIA
16	Test Data from Building and testing of CO2 expander and compressor	PDF report	ENOGIA
17	Test Data from Building and testing complete CO2 system (temperature, pressure, fluid flow, expander speed)	PDF report	UBRUN, ENOGIA, SYNESIS, EXP, CETRI
18	Environment data of testing site	Word, Excel, Designs	AME
19	Thermal data for the hot stream (surface temperature feed, temperature feed of air flow, coolant data)	Word, Excel, Designs	AME
20	Modelling data for FHPS (from various dates)	MS excel	UNBRUN, AME, ETHERM
21	Data from Manufacturing FHPS	Word, Excel, Designs	ETHERM
22	Data from flue gas thermal and chemical composition analyses (chemical composition and the dew point of each vapour) - Varius experimental campaigns on site (gas boiler, biscuit oven).	TXT File	UBRUN, ARLUY
23	Modelling data for design and optimisation of condensing economiser	Word, Excel, Designs	UBRUN
24	Testing data for condensing economiser (rate of corrosion)	Word, Excel, Designs	ETHERM
25	System Control Variables	Word, Excel, Designs	SYNESIS
26	Control and monitorin systems	Word, Excel, Designs	SYNESIS
27	Control system testing data	Word, Excel, Designs	ENOGIA
28	Database of papers and information related to coatings	Word, Excel	TEISTE
29	Testing data of coatings (corrosion, optical microscopy, scanning electron microscopy (SEM), and energy dispersive spectroscopy (EDS), AAS atomic absorption spectroscopy, UV/Vis spectroscopy, XRD)	Word and Excel files	AIT
30	Data from measuring water collection rate and heat transfer coefficient	Word and Excel files	TEISTE
30	Data from coating studies (microstructure, surface features by SEM, XRD etc. impact resistance, adherence and corrosion)	Word and Excel files	AIT
32	Performance evaluation data of TFC	Word, Excel, Designs	TEISTE
33	Performance evaluation data of sCO2	Word, Excel, Designs	ENOGIA, SYNESIS, UBRUN, EXP, AIT, TIESTE, CETRI
34	Mechanical & electrical design parameters of FHPS	Word, Excel, Designs	AME
35	Performance evaluation data of the Condensing Economiser	Word, Excel, Designs	UBRUN
36	Life cycle energy savings, life cycle cost and life cycle environmental performance of the technologies	Word, Excel, Designs	CUT, Tata Steel, Spirax, Etherm, AME, CETRI, Enogia, ARLUY