

Deliverable 6.2 Data Management Plan (DMP)

Document type **Deliverable D 6.2**

Document Version / Status **1.10 - First issue (living document)**

Primary Authors **Daria Barsuk, daria.barsuk@unilim.fr, UNILIM**

Distribution Level **PU (Public)**

Project Acronym **ATHOR**

Project Title **Advanced THERmomechanical multiscale mODelling of Refractory linings**

Grant Agreement Number **764987**

Project Website **www.etn-athor.eu**

Project Coordinator **Marc Huger, marc.huger@unilim.fr, UNILIM**

Document Contributors **Daria Barsuk, daria.barsuk@unilim.fr, UNILIM**
Marc Huger, marc.huger@unilim.fr, UNILIM
Eric Blond, eric.blond@univ-orleans.fr, UORL

Change log

History of changes				
Version	Date	Author (Organization)	Change	Page
1.00	11.04.2018	Daria Barsuk (UNILIM)	First version of the D 6.2 deliverable created	all
1.10	01.05.2018	Marc Huger (UNILIM)	Updated and ready for submission	all

Table of contents

- EXECUTIVE SUMMARY2
- 1. DATA SUMMARY2
- 2. FAIR DATA5
 - 2.1. Data storage5
 - 2.2. Making data findable, including provisions for metadata6
 - 2.3. Making data openly accessible.....6
 - 2.4. Making data interoperable.....7
 - 2.5. Increase data re-use (through clarifying licences)7
- 3. ALLOCATION OF RESOURCES7
- 4. DATA SECURITY8
- 5. ETHICAL ASPECTS.....8

Executive summary

As a reminder, a data management plan (DMP) is a formal document that outlines how data are to be handled both during a research project, and after the project is completed. The goal of a data management plan is to consider the many aspects of data management, metadata generation, data preservation, and analysis before the project begins; this ensures that data are well-managed in the present, and prepared for preservation in the future. Dissemination opportunities, targeted audiences and key messages will be identified and updated regularly. For the preparation of present version of this deliverable (at the beginning of ATHOR project) a number of reference data management documentation have been consulted by the authors and reviewed for the purpose to produce a self-consistent high-quality plan, based on best practices and experiences that authors consider related.

This deliverable provides the first version of the document on data management principles of the ETN ATHOR project. Present deliverable outlines how the research data collected or generated during the project will be handled by the consortium during and after the ATHOR action. It describes which standards and methodology for data collection and generation will be followed, and whether and how data will be shared and preserved. This document follows the template provided by the European Commission in the Participant Portal: Version 3.0, 26 July 2016.

The Data Management Plan is an evolving document of a public type, and will gain more precision and substance during the course of the project. It will be revised, if needed, in order to keep the information up to date.

1. Data Summary

In order to organise the best management of the data generated during ATHOR project, and as a check list during the data collection/generation process, we should keep in mind the following key points:

- What is the purpose of the data collection/generation in relation to the objectives of the project?
- What types and formats of data will the project generate/collect?
- Will you re-use any existing data and how?
- What is the origin of the data?
- What is the expected size of the data?
- To whom might it be useful ('data utility')?

The ATHOR project main objectives are directly linked to the problematics of the refractory materials for the iron and steel-making industry (I&S) - one of the most important sector of manufacturing and construction industry in Europe and in the world. A typical steel plant uses hundreds of types of refractories, each engineered for specific applications [American Ceramic Society Bulletin, Vol. 95, No.2]. It is therefore not surprising that refractories for I&S represent 60% of all refractory materials market today. While providing the top quality training in refractory materials and related fields for 15 young researchers and enabling close academia-industry cooperation, the principle objectives of this project are [Annex 1 of Grant Agreement 764987, part B page 7]:

- Explore the capability of advanced numerical modelling in order to design better materials and better refractory linings, consequently improving energy efficiency and thermomechanical properties of products
- Assist the European refractory and steelmaking industries to identify technological improvements by providing high performance characterization and modelling tools which can comfort their competitiveness
- Support the development of new characterization devices and modelling methods to offer solutions to current S/T challenges
- Assist the European refractory and steelmaking industries to identify technological improvements by providing high performance characterization and modelling tools, which can comfort their competitiveness.

Meeting these objectives requires cooperative work of all 15 ESRs together with their academic and industrial mentors. Each ESR will collect specific data, related to his PhD topic. As a result of exchanges within ATHOR consortium, it has become apparent to us that it is difficult to give an exact estimation of the kind and amount of data to be collected during the project. However, during the course of the project we plan to estimate the data, to be produced by each ESR, and produced as a result of other support activities such as project management and dissemination. Some preliminary results are accumulated in the table below.

Origin of the data			Purpose (in relation to the objectives of the project)	Type of data (formats), Data utility (Public or Not)
WP	Team	People		
Research activities: WP1, WP2, WP3, WP4	All consortium	ESR i (i = 1 to15)	<p>Analysed Data:</p> <ul style="list-style-type: none"> Analysed literature review Refractory materials properties analysis Quantitative model results 	MS Office/Open Office documents Public
		Supervisors (LSTC), WP1,2,3 4 Leaders with help of Project Manager	<p>Research Raw Data:</p> <ul style="list-style-type: none"> Literature review (compilation of documents) Industrial and models materials to be studied Experimental work description Refractory materials and subsystem characterisation In-situ temperature and strain fields measurements Models construction and boundary conditions 	MS Office/Open Office documents Just for ATHOR consortium
Training activities: WP5	Recruitment and Skill Progress Committee (RSPC)	RSPC Chair TKTC Chair WP5 Leader Lecturers	<p>Lectures, Workshop:</p> <ul style="list-style-type: none"> Powerpoints Video records E-Learning tools MOOC 	MS Office/Open Office documents Public
	Training and Knowledge Transfer Committee (TKTC)	Project Manager with help of ESR i (i = 1 to15)	<p>Training follow up:</p> <ul style="list-style-type: none"> PCDP documents Training plan Models from elsewhere 	MS Office/Open Office documents Just for ATHOR consortium
Dissemination activities: WP6	Training and Knowledge Transfer Committee (TKTC)	TKTC Chair WP6 Leader Project Manager	<p>Communication, Outreach:</p> <ul style="list-style-type: none"> Interactive exhibition Dedicated short videos Seminars / Newsletters Posters / Presentations Papers 	MS Office/Open Office documents Public
		with help of ESR i (i = 1 to15)	<p>Dissemination follow up:</p> <ul style="list-style-type: none"> Coming events listing Dissemination plan Models from elsewhere 	MS Office/Open Office documents Just for ATHOR consortium
Management activities WP7	Management Team (MT)	Project Coordinator Project Deputy Coordinator	<p>Project Description:</p> <ul style="list-style-type: none"> Objectives, Partners Work Packages, People ESRs and their Subjects Public Deliverables 	MS Office/Open Office documents Public
		Project Manager with help of ESR i (i = 1 to15)	<p>Project rules, follow up:</p> <ul style="list-style-type: none"> Consortium Agreement Financial statements Gantt, Actions plan Organisational aspect 	MS Office/Open Office documents Just for ATHOR consortium

Table 1. Estimated data that will be collected during the ATHOR project

Taking into account that first young researchers have joined the project about 4 months ago, we do not expect any considerable output in terms of their data provision plan at current stage. Therefore, precise information related to the data produced by ESRs will be given in the further versions of the DMP. The expected size can neither be predicted at this stage but it is reasonable to assume that it will reach the tens of Gigabyte range.

Main types of data to be generated in ATHOR can be approximately divided into 3 different groups:

- **Project rules and follow up data:** Grant and Consortium Agreements, Gantt Chart and Actions Plan, administrative and financial data, templates, surveys, management files
- **Research data:** it covers the data collected within the frames of the project subject including analysed data. In a research context, examples of such data include graphs and images, statistical data, parameters, experimental conditions, experimental observations, results of measurements etc. The focus will be made on the availability of this research data in digital form.
- **Data related to training activities:** Personal Career Development Plan (PCDP), Training & Visit Plan, Lecture notes, Powerpoints, Video Records, E-Learning tools, MOOC.
- **Data related to dissemination activities:** publications, presentations/posters, seminars/newsletters, dedicated short videos

These types of data include data of different confidentiality levels that can be schematically represented as in the Figure 1 (inspired by 5G! Pagoda D 1.2 - Open Data Management Plan, p.8). In this way, a dominant part of communication data and some part of research data is rendered public, while project data that ensures project's functioning is principally kept confidential (on UCloud).

The data generated by ESRs strongly depends on the individual doctoral projects, tools and research methods used within these projects. Whenever possible, the dataset will be made available online using the following formats:

- Text content: Acrobat PDF/A (.pdf); Comma-Separated Values (.csv), Microsoft Office or Open Office Formats (.docx, .xls, .pptx, .odt, .ods, .odp); Plain Text US-ASCII/ UTF-8 (.txt); XML (.xml)
- Graphic content (.jpg, .png, .svg .tif, .tiff)
- Audio content (.aif, .aiff, .wav)
- Video content (.avi, .mp4)
- Modeling data (.mat)

The project will assume the principle of using commonly used data formats for the reason of compatibility, efficiency and access.

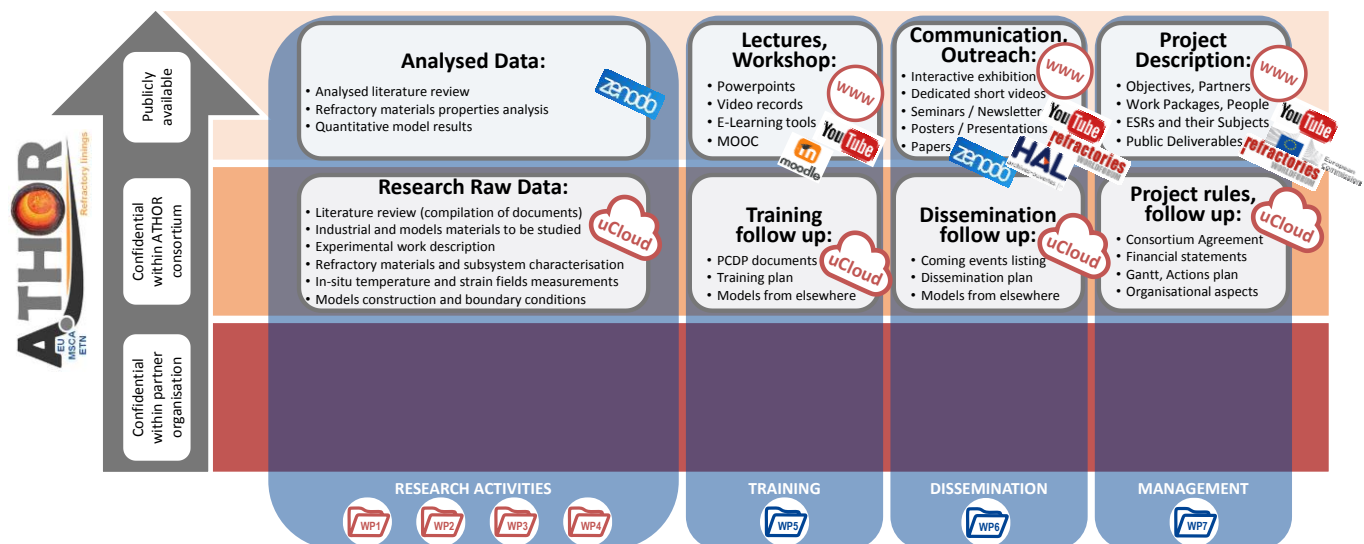


Figure 1. Distribution of ATHOR project data in the confidentiality grid

2. FAIR data

Findable, Accessible, Interoperable and Reusable

2.1. Data storage

The overall data produced and/or collected by each consortium member organisation has to be carefully stored and managed by this organisation. In a preliminary stage of production (and/or collection), a local storage by the authors is not excluded. When close to final version, all produced data have to be carefully stored by the authors in the central repository (<https://ucloud.unilim.fr>) dedicated to ATHOR project by the coordinating university of Limoges. All local and central repositories are to be secured using the latest security protocols. The access to the central repository is regulated by the project coordinator and project manager. It is provided for project consortium and to other linked parties upon request from a project team. The Ucloud platform is hosted by UNILIM servers, regular files backup is ensured by the local informatics services, additional archiving is made by designated project members on hard drive supports. The main Ucloud project repository is structured in the way, presented in the scheme of Figure 2.

Ucloud, being the central “data bank” feeds other platforms linked to the project. For instance, as it has been mentioned in Deliverable 7.2, at current stage ATHOR Principle Investigators with the lead of RWTH academic pole are establishing an e-learning program “Eleonor” in the frames of Modular Object-Oriented Dynamic Learning Environment (Moodle) that is planned to be hosted by UNILIM server. Such Moodle platform will store educational material on the subject of ATHOR project in written and visual form. It will also have two different access modes: private and public. The confidentiality status of each document deposited on the platform will be defined by IP owners of the document.

In the frames of project data management, all participants attempt to follow best practices for data generation, storage and sharing, i.e. document changelog, unified name attribution and appropriate repository are kept as clear as possible. The documents are preferably shared within the consortium via indication of its placement in a database. To facilitate document evaluation and review, all deliverables and official documents are created in agreement with established templates for main MS Office formats.

Each Work Package or task leader is responsible for timely preparation of corresponding deliverables and required materials, while the project coordinator assumes the responsibility for management activities and project administration.

ETN-ATHOR	
1	Proposal-Valuation-GA-CA
2	Potential Additional Funds
3	Team
4	Gantt Chart-Actions Plan
5	Meetings
6	Supervisory Board (SB)
7	Finance Committee (FC)
8	Industry Advisory Board (IAB)
9	Recruit. Skill Progress. Committee (RSPC)
10	ESR Council (ESRC)
11	Training & Knowledge Transfer Committee (TKTC)
12	State of the Art
13	WP1 - Improvement of measurements
14	WP2 - Advanced characterization
15	WP3 - Innovative modelling
16	WP4 - Advanced measurements
17	WP5 - Training, mobility
18	WP6 - Knowledge Dissemination
19	WP7 - Management Activities
20	ESRs working space
21	Dissemination
22	Scientific Publications
23	Deliverables and Milestones
24	Image gallery
25	Public

Figure 2. Directories tree of Ucloud repository of UNILIM

When a collection of data is ready to be published in public space, the last final version of these data currently stored on UCloud is upload on the most pertinent open access public platform. Depending of the type of data, it could be:

- **For project description:** ATHOR website (www.etn-athor.eu), Youtube Channel dedicate to ATHOR project, Refractories WorldForum (www.refractories-worldforum.com)
- **For research data:** Zenodo platform (<https://zenodo.org>)
- **For data related to training activities:** ATHOR website (www.etn-athor.eu), Youtube Channel dedicate to ATHOR project, Moodle platform dedicated to ATHOR project
- **For data related to dissemination activities:** ATHOR website (www.etn-athor.eu), Youtube Channel dedicate to ATHOR project, Refractories WorldForum (www.refractories-worldforum.com), CNRS Hal platform (<https://hal.archives-ouvertes.fr>), Zenodo platform (<https://zenodo.org>)

2.2. Making data findable, including provisions for metadata

In order to keep data findable, it is necessary to provide its metadata. Metadata is a systematic method for describing such resources and thereby improving access to them. Author, date created, date modified and file size are examples of very basic document metadata.

Considering the strongly interdisciplinary nature of the project, ATHOR's consortium favours the adoption of a broad and domain agnostic metadata standard that the EU recommends to its member states for recording information about research activity: the Common European Research Information Format (CERIF) standard is described at <http://www.eurocris.org/cerif/main-features-cerif>.

An additional advantage of a CERIF inspired standard is that ATHOR's DMP managing institution (University of Limoges) currently uses a research information system developed by Elsevier that implements the CERIF standard (PURE).

For publication data unique identifiers such as Digital Object Identifiers (DOI) will be used. According to authors knowledge it is the most common way for data identification. The repositories such as Zenodo or OpenAIRE (Open Access Infrastructure for Research in Europe), both or one of which is planned to be used for data publishing already provide persistent identifiers for data sets.

By one of the upcoming reporting periods, as soon as sufficient amount of data is produced within the project, consortium leaders will consider the distribution of survey template to all ESRs to collect the information on:

- Data set reference and name;
- Data set description;
- Data formats;
- Faced difficulties/risks while data collection and analysis;
- Standards and metadata
 - How is data created?
 - What standards or methodologies did you use?
 - How did you structure and name your folders and files?
 - How did you track the changelog?
- Data sharing;
- Archiving (storage and backup);
- Ethical issues;
- Other aspects (share of responsibilities within the team related to data lifecycle)

Currently, all data is stored on Ucloud platform and with the clear indication of data subject, authors and change log history when necessary.

2.3. Making data openly accessible

As mentioned in the previous section 2.3, after receiving the authorization of all concerned parties, it is planned to deposit the collected data on the most quoted repositories like [OpenAIRE](#) or [Zenodo](#) that allows researchers to deposit both publications and data, while providing tools to link them.

It is expected that data related to the social media, to any publicity, designated courses, open access publications, survey results, public deliverables will be made openly available by default. It is agreed within the consortium that the information, that has to be kept confidential within ATHOR will be marked with a special digital stamp with a mention "keep information private within ATHOR". Before any release of information, the authors of the document in question have to sign an "Authorization letter" clearly indicating his/her name, date, entity, the title of the document. A template of such authorization letter is available on Ucloud:

<https://ucloud.unilim.fr/public/authorization-letter-form>. This procedure was considered compulsory to avoid IP conflicts within the consortium and violation of the rules of good scientific practice and protection of personal data.

For some cases, in order to avoid multiplication of desynchronized versions of the data related to the same action and to moderate the use of virtual storage space, the consortium considers to preserve and make public only metadata, while removing raw data itself. The virtual address of the main and unique dataset (Ucloud platform) needs to be provided in parallel with metadata. In this way, the collected data remains findable and accessible.

Since the H2020 requirement for Open Access publishing is fully embraced by ATHOR project, the project will ensure both “green” (in addition to publication in subscription journals, the copy of an article is deposited into an institutional repository such as Research Repository UCD) and “gold” (publications available directly from the publisher after paying author’s fees, envisioned in the project’s budget) data publishing.

While ensuring the internal data storage and backup, the project obliges to publish the public results through the following channels:

- “Open data” section of the project website: www.etn-athor.eu menu.
- Zenodo (<https://zenodo.org>) central repository recommended by the Horizon2020 online manual where public deliverables and publications will be uploaded and connected to the OpenAIRE platform. The advantage of Zenodo is that it is “open in every sense”, hence, there is no need to explore any kind of arrangement with this repository, neither documentation, nor data access committee to access the data uploaded.

Since there is no any sufficient scientific data produced within the ATHOR project yet, these data repositories are currently empty.

- Diffuse or publish the appropriate type of data via ATHOR social channels, i.e. Facebook, LinkedIn, Youtube, Twitter.
- National portals for data publishing, for instance <http://theses.fr/> in France for depositing doctoral thesis manuscripts.

2.4. Making data interoperable

As mentioned in section 1. “Data Summary”, in order to comply with interoperability and re-usability requirements and to facilitate the exchange between researchers and institutions, best practices for file formats will be used in ATHOR project. When possible, data will be rendered available in the format consultable with the help of free of charge software (for example Open Office formats for text documents). The depositors will also strive for using a standard vocabulary for all data types present to allow interdisciplinary interoperability.

2.5. Increase data re-use (through clarifying licences)

It is possible to license a produced dataset. To do so, it will be necessary to attach Creative Commons Licence, according to the following guidelines <https://creativecommons.org/choose/> or <http://ufal.github.io/public-license-selector/> by integrating the appropriate abbreviation into the shared file. Since no research data has been produced to date, the specific question of its re-usability by third parties or usability period is not fully developed in the current version of the DMP document.

3. Allocation of resources

Generally, it is hard to predict the cost of data management activities, as many activities are an integral part of standard research activities and data analysis. Ideally, it is necessary to estimate the time or cost needed for activities related to data collection, data entry and transcription, data validation and documentation and the cost of preparing data for archiving and re-use. Those resources that include time and effort costs i.e. search costs, maintenance of technical infrastructure, individual preparation effort needed to use the infrastructure etc. are so-called non-monetary costs. Since ESRs and management team are the main producers of datasets in the project, all these costs are related to them. ATHOR consortium expects that monetary costs for FAIR data will be minor and will be mainly related to “gold” publishing of the articles, maintenance of hosting university servers for Ucloud and Moodle and engagement of external workforce for producing multimedia dissemination material.

Regarding the question of long term data preservation, no specific arrangements has been done in the consortium yet. However, with a great degree of confidence, it can be confirmed that it is the project coordinator with the help of local UNILM resources who will play the major role in this task.

4. Data security

The security of the central Ucloud <https://ucloud.unilim.fr/> repository and all other partner repositories is provided and guaranteed by the respective centres for information processing of these universities.

Access to the Ucloud database is managed by the coordinator and project manager. It is provided for project members and other parties upon request from a project team. This space is password protected and the security of this platform is guaranteed by the Informatics Systems Direction (DSI) of Limoges University. Regular back up of all data stored on UNILIM servers is ensured. The backup of Ucloud data is performed by the Limoges University server, the history of the content can be traced back to up to three months. The selected data, deposited in Ucloud space, will remain also available for 3 years after the end of the project.

In addition, if the project uses Zenodo repository for data sharing, its safety is guaranteed according to the product description (See <https://zenodo.org/features>)

5. Ethical aspects

According to the Annex 1 of Grant Agreement 764987 - Part B – p. 32, the ATHOR Consortium has taken into account all requested ethics issues. For example, the most common ethical issues include:

- the involvement of children, patients, vulnerable populations,
- the use of human embryonic stem cells,
- privacy and data protection issues,
- research on animals and non-human primates.

It also includes the avoidance of any breach of research integrity, which means, in particular, avoiding fabrication, falsification, plagiarism or other research misconduct.

More precisely, all the activities carried out under the ATHOR project comply with ethical principles and relevant national, EU and international legislation, for example the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights. The tasks for ATHOR only concern basic research activities and the project does not involve humans, animals or cells. Due to the fact that the main domain of the ATHOR project activity is related to materials science with the focus on refractory materials, the risk of having ethics issues during the project is extremely limited. Either way, within the ATHOR DoA Part A, the workpackage 8 is devoted to the ethics issues which sets out the 'ethics requirements' that the ATHOR project must comply with. One deliverable will be provided: D8.1 NEC - Requirement No. 1. In the framework of D8.1, all beneficiaries and partner organisations must confirm that the ethical standards and guidelines of Horizon2020 will be rigorously applied, regardless of the country in which the research is carried out.

ATHOR's partners are not planning to use any harmful material, or process which likely emits harmful materials. They do not use elements that may cause harm to the environment, to animals or plants. In any case, all the partners will follow their internal protocols to treat any material according to the national law and EU legislation. In that way, all chemical waste is collected and processed by a central university facility in the Universities involved within the ATHOR project. All wastes are recycled or appropriately deposited. Moreover, their respective researches do not deal with endangered fauna and/or flora /protected areas. No tests on humans or animals are planned.

ATHOR'S partners will not use nano-material in their research and they do not do harm to the environment.