



DELIVERABLE D7.4

First RPX update of the Data Management Plan

WP7

Lead Beneficiary: 1-UNIPD

Author(s): Eloisa Di Sipio¹, L. Pockele², A. Romanowski³, O. Stenmeier⁴, R. Pasquali⁵, A. Manzella⁶

Author(s) affiliations: ¹UNIPD, ²RED, ³Prevent, ⁴FhG-IAPT, ⁵GEOSERV, ⁶CNR

Date: 28/02/2023

Dissemination Level

PU	Public, fully open	X
SEN	Sensitive - limited under the conditions of the Grant Agreement	
CI	EU classified - RESTREINT-UE/EU-RESTRICTED, CONFIDENTIEL-UE/EU-CONFIDENTIAL, SECRET-UE/EU-SECRET under Decision 2015/444	



This research is funded by the European Union (G.A. 101046937). The views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EISMEA. Neither the European Union nor the granting authority can be held responsible for them.

Document History

Version	Date	Authors	Description
1	15/01/2023	E. Di Sipio (UNIPD)	Creation of the document
2	08/02/2023	E. Di Sipio (UNIPD), R. Pasquali (Geoserv), L. Pockelè (RED)	First Draft ready
3	10/02/2023	A. Romanowsky (Prevent)	Review
5	15/02/2023	A.Manzella, G. Montegrossi (CNR)	Review
6	17/02/2023	E. Di Sipio (UNIPD)	Second draft ready
7	27/02/2023	O. Stenmeier (FhG-IAPT)	Review
8	28/02/2023	E. Di Sipio (UNIPD)	Final Version for Upload to SyGMA

Disclaimer

This document is the property of the DeepU Consortium.

This document may not be copied, reproduced, or modified in the whole or in the part for any purpose without written permission from the DeepU Coordinator with acceptance of the Project Consortium.

This publication was completed with the support of the European Innovation Council and SMEs Executive Agency (EISMEA) under the HORIZON-EIC-2021-PATHFINDEROPEN-01 programme. This research is funded by the European Union (G.A. 101046937). However, the views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or EISMEA. Neither the European Union nor the granting authority can be held responsible for them.

TABLE OF CONTENTS

PUBLISHABLE SUMMARY	4
ABBREVIATIONS	5
1. INTRODUCTION	6
2. DATA SUMMARY	7
3. FAIR DATA	10
3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA.....	10
3.2 MAKING DATA ACCESSIBLE	11
3.2.1 <i>Repository</i>	11
3.2.2 <i>Data</i>	12
3.2.3 <i>Metadata</i>	13
3.3 MAKING DATA INTEROPERABLE	13
3.4 INCREASE DATA RE-USE.....	14
4. OTHER RESEARCH OUTPUTS.....	16
5. ALLOCATION OF RESOURCES	16
6. DATA SECURITY.....	17
7. ETHICS.....	18
8. OTHER ISSUES	18
8. REFERENCES	19
APPENDIX A.....	20
QUESTIONNAIRES ON RESEARCH DATA.....	20
WP1 - PREVENT	21
WP2 – FHG-IAPT	24
WP3 - UNIPD	27
WP4 – GEOSERV.....	31
WP5 – RED	34
APPENDIX B.....	37
GUIDELINES TO UPLOAD DATASETS ON THE ZENODO REPOSITORY	37

Publishable summary

The D7.4 “First RPX update of the Data management plan” is a public document delivered in the context of WP7, Task 7.1: General project direction, management and communication with regard to data management, collection, generation, storage and preservation.

This document describes the update to the Data Management Plan (DMP, D7.3) for “Deep U-tube heat exchanger breakthrough: combining laser and cryogenic gas for geothermal energy exploitation (DeepU)” Project.

The use of a Data Management Plan is required for all participating projects supported by the European Innovation Council and SMEs Executive Agency (EISMEA) under the HORIZON-EIC-2021-PATHFINDEROPEN-01 programme.

This document details the type of data produced or collected during the project activity according to each WP. Information about (the kind of data collected, processed and/or generated, the methodology & standards to be applied for their preservation, the handling and cure of research data (shared/made open access) are included.

This deliverable represents the DMP at the present stage of the project: during the project life cycle, new information will update the DMP. In fact, the datasets here proposed could evolve during the lifespan of the project, particularly whenever significant changes arise such as changes in Consortium policies. Future modifications will be tracked during each periodic report. In detail, a final DPM update is foreseen at month 36 (D7.8).

Abbreviations

CA	Consortium Agreement
DeepU	Deep U-tube heat exchanger breakthrough: combining laser and cryogenic gas for geothermal energy exploitation
DMP	Data Management Plan
DoA	Description of Action
DOI	Digital Object Identifier
EC	European Commission
FAIR	findable, accessible, interoperable and reusable
GA	Grant Agreement
IPR	Intellectual Property Rights
OA	Open Access
ORD	Open Research Data
PC	Project Coordinator
PC-Co	Project Co-coordinator
RD	Research Data
SC	Steering Committee
T	Task
WP	Work Package

1. INTRODUCTION

A Data Management Plan (DMP) has been developed using FAIR data principles (D7.3).

DeepU consortium adopt **Open Science** practices to increase the chances of delivering the results to a wide community of researchers. Whilst respecting the IPR policy, the data/research outputs will be made findable, accessible, interoperable and reusable (FAIR).

The DMP outlines what datasets the project will generate and compile, and how these datasets will be made accessible and stored. The DMP also describes measures that have been taken to safeguard and protect sensitive data and emphasizes that the produced results must be easily located and accessible.

At present, after one year of activity, the first data collected by the project must be organized. The research data generally consist in datasets of parameters, generated in the action by the Consortium according to the different WPs. These data will be used during the life span of the project, examined and compared between different WPs, and to serve as a basis for reasoning, discussion and/or calculation.

Open access to research data consists of the right to access and reuse digital research data under the terms and conditions set out in the DeepU Grant Agreement.

Even though the Consortium greatly endorse the precepts of open science, nevertheless, data sharing in the open domain can be restricted as a legitimate reason to protect results that can reasonably be expected (e.g. commercial or industrial exploitation). Strategies to limit such restrictions could include anonymizing or aggregating data, agreeing on a limited embargo period or publishing selected datasets not completely open.

Most of the WPs and the relative deliverables indirectly influence this document, due to the data they contain. For this reason, this document is an update of the DeepU Data Management Plan (D7.3). In detail, the data management of the preliminary data collected or expected to be collected in each project WP is addressed, remaining valid the general conditions already stated in D7.3.

The results of the “Questionnaire on Research Data” template, distributed to the whole consortium, are here presented and discussed. In fact, the DeepU DMP is intended to be a ‘living’ document, where for each WP the expected datasets are described, highlighting the possible restriction in the use and availability of data due mainly to IPR and patent liability.

In the future, the DMP will be reviewed and updated taking into consideration the project progress, both in terms of science and innovation, considering different data sharing strategies following the principle ‘as open as possible as closed as necessary’, respecting the beneficiary’s legitimate interests, regarding commercial exploitation and patentability.

UNIPD is responsible for disseminating this DMP to all project partners. Each project partner is responsible for managing their data, metadata, and insuring their data meets the quality standard set out in the DeepU Quality Assurance Plan (D7.1) and in the Data Management Plan (7.3).

2. DATA SUMMARY

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

During the project life cycle, new data are created and expected datasets are expanded with new information. However, in each WP the re-use of any existing data has been considered, has reported in the questionnaire on Research Data (Appendix A) shared with partners at the first GA (6M).

Hereafter, divided by WP, a summary of the main existing data reuse is shown.

Table 1: Overview of the existing data re-use in each WP.

WP	Existing database	Kind of data
WP1	<ul style="list-style-type: none"> - Internal data generated by WP1 - existing patent reviews including already filed patents of DeepU partner 	<ul style="list-style-type: none"> - technological data - technical drawings
WP2	<ul style="list-style-type: none"> - Internal data generated by WP1 - Internal data generated by WP2 - Existing computational design for laser-optical components 	<ul style="list-style-type: none"> - Technical drawings and CAD models - Process parameters - Measurement data of the environmental conditions
WP3	<ul style="list-style-type: none"> - P³ PetroPhysical Property - Geo4civhic Project thermal properties - Previous studies on melted rocks - Internal data generated by WP1 and WP2 	<ul style="list-style-type: none"> - Rocks mineralogical and petrographic data - Rocks thermo-physical properties - Data about laser technology and cryogenic gas usage
WP4	<ul style="list-style-type: none"> - Internal data generated by WP1 and WP2 - Internal data generated by WP3 	<ul style="list-style-type: none"> - Data about laser technology and cryogenic gas usage - Information on rock/laser interaction and vitrification
WP5	<ul style="list-style-type: none"> - existing data on drilling costs from literature and drilling operators - previous studies on power plants modeling and deep geothermal energy extraction - existing patent reviews including already filed patents of DeepU partner 	<ul style="list-style-type: none"> - Economic data - Energy data - Information data

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

Table 2 provides a preliminary summary of the Data type and the corresponding Datasets that are expected to be produced by each project WP. This brief outline of the datasets gives an instant overview on the expected basic information of the RD.

Table 2: Data type and Datasets that are expected to be produced by each project WP

WPs	Product	File formats
WP1	Text Design	PDF STEP

WP2	Text Models, Images Spreadsheets	Plain text, PDF STEP, STL, JPEG, TIFF CSV, Excel
WP3	Text Spreadsheets Images	Plain text, XML, HTML, PDF CSV, Tab delimited values, PDF, Excel JPEG, TIFF; PNG, SVG
WP4	Text Spreadsheets	Plain text, XML, HTML, PDF CSV, Tab delimited values, PDF, Excel
WP5	Text Spreadsheets	Plain text, XML, HTML, PDF CSV, Tab delimited values, PDF, Excel, numerical data to be organized

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

The DeepU Data Management Plan (DMP) aims to provide a strategy for managing data generated and collected during the project and to optimize access to and re-use of research data. Data generated during the project can be divided into the following groups:

- Data collected from the outside of the project: used or modified data acquired / collected from existing databases, repositories, open publications because functional for the project activities planned in the different WPs.
- Original data generated by project partners during the laboratory and technological activities. These datasets consist of compiling raw data provided by partners working in the different WPs in the format of agreed metrics.

The main reasons for the re-use of existing datasets are summarized in Table 3.

Table 3: Reasons for the re-use of data in each WP.

WP	Reason for re-use
WP1	To manufacture new laser drill rods and to consider new materials to improve their performance
WP2	Manufacture of spare parts, repetition of process settings, traceability of the boundary conditions during the process
WP3	Reference data considered for DeepU rock samples petrophysical analysis and interpretation
WP4	To review published regulations and standards to develop recommendations on technology operational options and Environmental Health and Safety (EHS) conditions
WP5	Assess indicative cost reduction potential of DeepU drilling technology Assess cost of renewable electricity and heat extraction potential from DeepU solution by modelling Review IPR potential of DeepU technologies

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

The data generated through measurements and laboratory activities in each WP are expected to have a maximum size of 10 MB for each dataset. Open access publication on peer-review journal will respect this size limit.

Anyway, the maximum file size limit of each dataset is expected to be lower than 50 GB, that is the maximum file size allowed by the repository selected by DeepU project and described in section 2.2 of D7.3.

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

Data collected during DeepU activities concern mainly laboratory and technological activities. Data from existing standards (CEN, ISO) and best practice guidelines on the project innovations will be gathered and assessed on the basis of the replicability of the project technologies to the drilling and geothermal industries (Table 4).

Table 4: Origin/provenance of data in each WP.

WP	Origin provenance of the data
WP1	Internal data generated by WP1, existing patent reviews including already filed patents of DeepU partner, previous knowledge of the beneficiary.
WP2	Internal project data from WP1 and WP2; existing calculation models for the design of laser-optical components
WP3	Internal project data from WP1, WP2, WP3; scientific publications, national and international standard bodies.
WP4	Government publications, national and International standard bodies, internal project deliverables from WP1, WP2, WP3 and WP4
WP5	Scientific literature, deep drilling operators, power plant suppliers, patent literature, modelling performed in the frame of DeepU Project

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

The data generated in the project will be very beneficial to a variety of stakeholders including: researchers, policy makers, public funders, device developers, utilities, private investors. For each WP a list of specific stakeholders is shown in Table 5. However, in accordance with Annex 5 of the grant agreement, the project partners are obliged to protect the results where these can be expected to be commercially or industrially exploited.

Table 5: WP data utility.

WP	Data utility
WP1	Future stakeholders, investors, manufacturers
WP2	Project partners, future stakeholders, public authorities
WP3	Project partners (for checking drilling effects on rocks and as input to model simulation of underground closed loop heat extraction), future stakeholders, researchers, geothermal and oil and gas companies, public authorities, general public.

WP4	Project partners, public authorities.
WP5	Project partners, future stakeholders, i.e. drilling contractors and multi-utility companies, oil and gas companies, ventures capitalists, public authorities, and geothermal sector for defining economics of the proposed solution.

3. FAIR data

The DeepU DMP (D7.3) applies the Findable, Accessible, Interoperable, Reusable (FAIR) approach for the project's results. This section of D7.4 takes up Chapter 3 of D7.3 by inserting some clarifications.

3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Will data be identified by a persistent identifier?

As stated in D7.3, with the exception of RD that are not classifiable as ORD due to commercial exploitability reasons, the RD will be always included in the open access repository and they will be identified with a unique reference code (DOI), despite their exact shareability level. Shareability level and conditions may vary during the project: any changes will be operated in the data repository system as a consequence.

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.

Metadata has value to both the original creator of a data set and other potential users. Complete metadata allows researchers to locate data they created and recall the circumstances and context under which they created and analyzed the data. It allows researchers outside of the original research team to discover, understand and use the data.

Therefore, all DeepU documents will be identifiable based on a common naming convention. To ensure document and data control, each document and data set shall be uniquely identifiable. Each deliverable and data set must be associated unique document name to ensure version control. The deliverable and data identifier must be used in the deliverable filename.

The data identifier for the deliverable must be: <Deliverable identifier><Up-to-five-words-from the data name>_<project name acronym><followed by the version number (v01, v02, def)>

Example: D7.1_QualityAssurancePlan_DeepU_def

Metadata will be associated to each dataset following the metadata standards directives (<http://rd-alliance.github.io/metadata-directory/>, <http://jennriley.com/metadata/>, <https://www.dcc.ac.uk/guidance/standards/metadata>).

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

Keywords will be provided in the metadata in order to clearly identify the type of data contained in the report and increase the possibility for discovery and potential re-use (D7.3).

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

Metadata Harvesting is favoured by the use of Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), a low-barrier mechanism for repository interoperability. In addition, as described in detail in D7.3, some of the metadata is used to index the records to make retrieval easier.

3.2 MAKING DATA ACCESSIBLE

3.2.1 Repository

Having in mind that the "Guidelines for FAIR Data Management in Horizon 2020" [1] recognize the need to balance openness and protection of scientific information, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions, DeepU Consortium has selected **the Zenodo repository** (<https://zenodo.org>) [2] for storing and preserving the project data. A reference to the DeepU project will be attributed to each uploaded dataset, and the same will be done for the new data that will be produced and organized as coherent datasets. Zenodo has been used in other European projects, and is the backend of the the **European Geothermal RD&I Document Search Engine (EGRISE)** [3]. EGRISE was developed in 2018 in the frame of the European technologies and Innovation Platform on Deep Geothermal (ETIP-DG), with the aim of organizing a catalog of the European Projects' public results referring to deep geothermal topics. A "Deep Geothermal" community has been established within Zenodo and the content is growing with time.

Will the data be deposited in a trusted repository?

Zenodo is hosted by CERN and is funded, among others, by the European Commission via the OpenAIRE projects

- FP7: OpenAIRE (246686), OpenAIREplus (283595)
- Horizon 2020: OpenAIRE2020 (643410), OpenAIRE-Connect (731011), OpenAIRE-Advance (777541), and OpenAIRE-Nexus (101017452).

Zenodo is developed and supported as a marginal activity, and hosted on top of existing infrastructure and services at CERN, in order to reduce operational costs and rely on existing efforts for High Energy Physics. CERN has some of the world's top experts in running large scale research data infrastructures and digital repositories in order to deliver a trusted digital repository.

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

The Zenodo service ("Zenodo") is offered by CERN as part of its mission to make available the results of its work.

Access to Zenodo's content is open to all, for non-military purposes only. Content may be uploaded free of charge by those without ready access to an organized data center.

The terms of use are available at the following link <https://about.zenodo.org/terms/>.

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

Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citeable. Zenodo further supports harvesting of all content via the OAI-PMH protocol. For further information, please check the "DOI versioning" at the following link <https://help.zenodo.org/>.

3.2.2 Data

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

A mixed situation exists in terms of dataset openness. Although all datasets are shared within the Consortium, only a part of that amount is labelled as ORD at this moment. Furthermore, data sharing related to the technological innovation in WP1 and WP2 will be restricted as a legitimate reason to protect results that can reasonably be expected to be commercially or industrially exploited. Strategies to reduce such restriction will include agreeing on a limited embargo period and will be the focus also of WP5.

In the Questionnaire on Research Data (Appendix A), each WP leader stated the status of restrictions related to the datasets now being collected and defined. Anyway, the DeepU consortium shareability policy always aims for the maximum openness and reusability of data.

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.

Materials generated under the project will be disseminated in accordance with partners policies.

Publication of data shall occur during the project, if appropriate, or at the end of the project, consistent with normal scientific practices. Research data which documents, supports and validates research findings will be made available after the main findings from the final research dataset have been accepted for publication and/or 6 months after the laboratory measurements, the field data collection and numerical simulation are completed, according to the project workplan.

When an embargo status is defined, the partners will provide an end date for the embargo. The repository will restrict access to the data until the end of the embargo period; at which time, the content will become publicly available automatically.

Data sensible for patent requirements, as those related to the laser drill rod, will be subjected to access restriction, respecting the beneficiary's legitimate interests, including those regarding commercial exploitation. A patent strategy will be defined under WP5 and the access to restricted/sensitive data will be regulated considering also license agreements with third parties outside the project. Time needed to grant patent and to secure economic exploitation rights will be considered in the IPR policy.

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

Access to databases and associated software tools generated under the project will be available for educational, research and non-profit purposes related to geothermal energy developments. Such access will be provided using web-based applications, as appropriate. ZENODO digital repository tool, a simple and innovative service that enables researchers, scientists, EU projects and institutions to share, preserve and showcase multidisciplinary research results (data and publications) will be used as storage system for DeepU Project.

Access to data objects: Files may be deposited under closed, open, or embargoed access. Files deposited under closed access are protected against unauthorized access at all levels. Access to

metadata and data files is provided over standard protocols such as HTTP and OAI-PMH. Use and re-use is subject to the license under which the data objects were deposited.

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

Restricted Access is also foreseen for data containing potential for patentability (see CA and GA agreement). These files will not be made publicly available and sharing will be made possible only by the approval of the creator of the original file.

The metadata about the restricted data will be shared in a repository in order to increase findability/discoverability and explain to others what they have to do to request access.

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

The identity of the project partner accessing the data will be ascertained via website login through the private section.

Access to ORD files deposited under OA conditions is free and will follow the Zenodo protocol.

The access to restricted/sensitive data will be regulated by the SC and it is subjected by the approval of the creator of the original file. A direct request must be sent to the Consortium, so the identity of the person accessing these data can be ascertained.

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

The Steering Committee meeting is in charge to determine the data access policy of DeepU project. Every 6 months, during the General Assembly, the data access policy could be reviewed.

3.2.3 Metadata

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

Metadata is licensed under CC0, except for email addresses. All metadata is exported via OAI-PMH and can be harvested. The metadata format is described in section 2.3.

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

The research is stored safely for the future in CERN's Data Centre for as long as CERN exists. Therefore, data and metadata are guaranteed to remain available after the end of the project.

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)?

When data will be prepared in a format different from the standards described in Table 2, the creator of the data will be responsible to provide reference to the software needed to access or read data and, in case, to include the relevant open source code, if available. More detail will be provided by partners once the questionnaire in Appendix A will be updated.

3.3 MAKING DATA INTEROPERABLE

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across

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

Each dataset will present data collected and shared in a standardised way using a standard format for that data type (see Table 2). References to any software required to run and read the data will be provided. Given the scope of this project, publicly available software the data will be readable by publicly available softwares. Barriers to access through interoperability issues are not anticipated.

The metadata format will follow the convention of the hosting research data repository. Instruction for uploading in Zenodo and inserting metadata is set out in Appendix B.

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?

The use of uncommon or specific ontologies or vocabularies is not foreseen. In case of uncommon ontologies generation, vocabularies and ontologies to facilitate the data redeang and reuse will be provided.

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

Any reference to other data and datasets from previous research will be clearly stated.

3.4 INCREASE DATA RE-USE

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

Data used for validate scientific open access publication will be available for data generation and validation/interpretation /re-use.

There are different ways to validate data analysis and facilitate data re-use. In detail, DeepU project is expected to make use of:

- 'readme' file: any information that cannot be recorded in a structured way (i.e. as the values of fields in a data or metadata file) are recorded as free text within a readme file.
- file formats within the data file: it is possible to record information in addition to the main data content for example in XML or CSV standard files, that provide a way of recording sampling strategies and procedures as well as measurement values, variables and units of measurements.
- separate metadata file: some disciplines have developed special file formats or data structures for recording supporting information.
- published journal article: some of the information needed to understand data would normally be provided in a journal article reporting the research. In order to prevent duplication of effort, it is possible to refer to an article to provide more information about a dataset, but before doing so you should be sure that (a) the article provides sufficient detail and (b) that the article will be available as open access.

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

The digital research data generated in the action ('data'), in line with the FAIR principles, as soon as possible will be deposited in a trusted repository. Open access to the deposited data will be ensured via the repository, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC 0) or a licence with equivalent rights, following the principle 'as open as possible as closed as necessary', respecting the beneficiary's legitimate interests, including regarding commercial exploitation, or any other constraints, justifying it in the DMP.

Metadata of deposited publications will be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (in particular machine actionable) and provide information at least about the following: publication (author(s), title, date of publication, publication venue); Horizon Europe funding; grant project name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action and, if possible, for their organizations and the grant. Where applicable, the metadata must include persistent identifiers for any research output or any other tools and instruments needed to validate the conclusions of the publication.

Open access to peer-reviewed scientific publications states that at the latest at the time of publication, a machine-readable electronic copy of the published version or the final peer-reviewed manuscript accepted for publication, is deposited in a trusted repository for scientific publications. In addition, immediate open access is provided to the deposited publication via the repository, under the latest available version of the Creative Commons Attribution International Public Licence (CC BY) or a licence with equivalent rights. The licence may exclude commercial uses and derivative works (e.g. CC BY-NC, CC BY-ND).

License agreement will be approved in WP5 in case of data sensible for patent request and commercial exploitation. In this case, no data re-use will be allowed outside the consortium until patent requests are publicly available.

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

Even though the Consortium greatly endorse the precepts of open science, nevertheless, data sharing in the open domain can be restricted as a legitimate reason to protect results that can reasonably be expected (e.g. commercial or industrial exploitation). Strategies to limit such restrictions could include anonymizing or aggregating data, agreeing on a limited embargo period or publishing selected datasets not completely open.

When no restrictions are specified and justified, the data could be used by third parties, in particular after the end of the project. The data will remain re-usable at undetermined time

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

Data provenance is the documentation of where a piece of data comes from and the processes and methodology by which it was produced. Put simply, provenance answers the questions of why and how the data was produced, where, when and by whom. This kind of information will be recorded as metadata to confirm the authenticity of data and to enable it to be reused.

In its simplest form, provenance can be recorded in a single README text file that describes the data collection and processing methods used. Provenance can also be recorded in a more structured way using specific elements in very generic metadata standards such as Dublin Core, to discipline-specific metadata standards such as ISO 19115-2.

Are all relevant data quality assurance processes well described?

Data quality assurance is the process of identifying and eliminating anomalies by means of data profiling and cleansing. Data quality control will be performed by the partner responsible for the RD creation. Duplicates, outliers, errors, and missing information will be detected in order to obtain accurate, complete, and consistent data, essential to track the progress of current projects and proposed initiatives. The following practices will be applied to each RD creation:

- **Relevance:** the data should be interpretable. This means that the partner has appropriate data processing methods, that the data format is interpretable by the partner software and that the legal conditions allow the partner to use such data.
- **Accuracy:** Ensuring the accuracy of the data by techniques like data filtering and outlier detection.
- **Consistency of data:** By checking internal and external validity of the data consistency can be ensured.
- **Timeliness:** The more up to date data suggests more precise calculations / organization.
- **Compliance:** it is important to check whether the data used complies with DeepU project progress or not.

4. OTHER RESEARCH OUTPUTS

This section of D7.4 takes up Chapter 4 of D7.3 with a few changes.

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

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

In case of digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.) project outputs, the partner responsible for it (i.e. Prevent for D1.3 Drill strings prototypes) provided a plan for its management answering to the Questionnaire on research data (Appendix A). The beneficiary's legitimate interests, including regarding commercial exploitation or any other constraints, will be respected.

5. ALLOCATION OF RESOURCES

This section of D7.4 takes up Chapter 5 of D7.3 with a few changes.

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 activities related to making the data/outputs open access are anticipated to be covered within the allocated budget for each work package. Further investigation of potential cost related to a repository need to be done. The repository will ensure that data is stored safely and securely and in full compliance with European Union data protection laws and in accordance with Annex 5 (Article 16-17) of GA.

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)

The costs related to OA to research data are eligible as part of the DeepU Horizon Europe grant. The costs of making scientific publications, hosting a project website and the partners and open access data repositories are contained within the DeepU budget as eligible costs.

Who will be responsible for data management in your project?

UNIPD as Project Coordinator, in collaboration with consortium partners and especially CNR (WP6 leader), is responsible to manage the data generated during the project. The PC identified an appropriate data repository to store and safeguard the datasets able to ensure that data are readily accessible. All WPs leaders will be responsible of the quality of data generated during the project and will assist the PC in organizing the different datasets.

CNR, in charge of WP6: Dissemination and Communication, in collaboration with consortium partners, will oversee the identification of which datasets will be disseminated and the most appropriate means of disseminating these data.

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)?

Resources for long term preservation, associated costs and potential value, as well as how data will be kept beyond the project and how long, are discussed by the Consortium's General Assembly (GA) at the DeepU Steering Committee (SC) meeting. The use of Zenodo repository is useful in dealing with the long-term preservation of data related to DeepU project.

The PC is responsible for all data management during and after data collection, while each WP leader is responsible for preparing the datasets related to the corresponding WP, indicating which data can be made public and which are bound by IPR agreements. Copies of all datasets will be saved by WP leaders to local servers and external drives, while the final version will be updated in Zenodo.

6. DATA SECURITY

This section of D7.4 takes up Chapter 6 of D7.3 with a few changes.

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)?

According to the best practices in the matter of data preservation, all files uploaded to Zenodo will be stored in CERN's 18 petabytes disk cluster. In that infrastructure, each file copy has two replicas located on different disk servers. For each file, two independent MD5 checksums are stored. One checksum is stored by Invenio (a Free Open Source Software providing repository/document management platform, an integrated library system and a code library to build large-scale information systems) and used to detect changes to files made from outside of Invenio. The other checksum is stored by EOS (the primary low latency storage infrastructure for physics data from the Large Hadron Collider), and used for automatic detection and recovery of file corruption on disks.

From the security standpoint, physical disks in which uploaded data will be stored are set at CERN Data Centre located on CERN premises, all physical accesses will be restricted to a limited number of staffs with appropriate training, and who have been granted access in line with their professional duties. Remote access to the servers is restricted to Zenodo staff and the operating system and

installed applications are kept updated with latest security patches via the automatic configuration management system Puppet. CERN Security Team runs both hosts, network-based intrusion detection systems, and monitors the traffic flow, pattern and contents into and out of CERN networks in order to detect attacks. All access to zenodo.org happens over HTTPS, except for static documentation pages which are hosted on GitHub Pages.

Zenodo stores user passwords using strong cryptographic password hashing algorithms. Users' access tokens to GitHub [4] and ORCID [5] are stored encrypted too.

Zenodo allows users to upload files under closed access. Closed access means that zenodo.org users will not be able to access the files you uploaded. The files are however stored unencrypted and may be viewed by Zenodo operational staff under specific conditions. This means that "closed access" on Zenodo is not suitable for secret or confidential data.

In case of sensitive "closed access", "secret" or "strictly confidential" datasets, data owner and the involved DeepU partners will evaluate together the most suitable ad hoc solution to protect and secure such information, including hosting in DeepU website [6].

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

The use of Zenodo repository comply with the request of long-term preservation and curation of project data, ensuring also the data security over time.

Using standard or open data formats ensures longer-term usability of data. Therefore, project files will be converted into standard open formats for long term preservation of the data.

7. ETHICS

A dedicated section (section 4) has been included in the DoA of DeepU project to ensure that ethical principles are used in the framework of the project, including among others, specific aspects about data collecting in real cases and scientific moral belief.

There are no known ethical or legal issues that can have an impact on data sharing, because research activities within the project do not directly involve the collection of any personal data or other sensitive information. Furthermore, metadata too are not relevant from this standpoint, as the data collection process will include anonymizing or aggregating data strategies, whenever such actions will be necessary.

In the whole processes of data collection, data analysis, dataset creation, dataset storing and dataset access, the Consortium will make certain that the General Data Protection Regulation (GDPR), which has entered into force in May 2018, is ensured, especially in regards to protection of private personal data.

8. 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)?

The lead partner and all beneficiaries are subject to the CA and GA. All data must be collected, stored and disseminated in accordance with these Agreements.

8. REFERENCES

- [1] H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020 (2016). EUROPEAN COMMISSION Directorate-General for Research & Innovation.
- [2] <http://zenodo.org/> - ZENODO data repository website.
- [3] Trumpy, E., Gola, G., Botteghi, S., Pellizzone, A., Sorin, P., Dumas, P., Pinzuti, V., Laenen, B., Manzella, A. THE RD&I DOCUMENT SEARCH ENGINE OF ETIP-DG. European Geothermal Congress 2019, Den Haag, The Netherlands, 11-14 June 2019
- [4] <https://github.com/> - GitHub Inc. web-based hosting service for computer code.
- [5] <https://orcid.org/> - ORCID website.
- [6] DeepU EU Project: <https://www.deepu.eu/>

Appendix A

QUESTIONNAIRES ON RESEARCH DATA

The Questionnaire on Research Data filled in by each WP leader and summarized in D7.4 are shown hereafter.

The answers were provided by each WP leader and concern an overview of datasets management for each WP.

WP1 - Prevent

TOPIC Q. No

General information	A1	Please, enter your DeepJ partner name	Prevent GmbH
	A2	At this time, did your actions during the project generated any research data?	Yes
	A3	At this time, did you used or modified any data acquired from the outside of the project?	Yes
	A4	At this time, did you published any open access scientific peer-reviewed publication that includes research data?	No
	A5	At this time, how many dataset have you generated, used or modified?	None
	A6	Do you plan to generate even more research data during the lifespan of the project?	Yes
	A7	Do you plan to use or modify data acquired from the outside of the project during the lifespan of the project?	No
	A8	Do you plan to publish open access to sceintific peer-reviewed publications that make use of research data during the lifespan of the project?	No
	A9	How many dataset do you plan to generate, use or modify during the whole lifespan of the project?	5 datasets

TOPIC Q. No

Data summary	B1	What is the short name that you would use to describe this specific dataset?	laser drill rods drawing
	B2	What is the purpose of the data collection/generation and its relation to the tives of the project?	without these drawings noone can drill with a laser beam
	B3	What type and formats of data will the WP generate/collect?	PDF-files and STEP-files (STEP means „Standard for the Exchange of Product model data“ according to ISO Norm 10303)
	B4	Will you re-use any existing data and how?	yes, we will reuse existing data and modify them and find new laser drill rods and new materials to reach a better performance
	B5	What is the origin of the data?	previous knowledge and patent of the beneficiary
	B6	What is the expected size of the data?	design drawings < 10 MB
	B7	To whom might it be useful ('data utility')?	for every Investor or manufacturer who wants to manufacture laser drill rods

TOPIC	Q. No		
Data summary	C1	Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.	Data sensible for patent requirements, as those related to the laser drill rod, will be subjected to access restriction, respecting the beneficiary's legitimate interests, including those regarding commercial exploitation. A patent strategy will be defined under WP5 and the access to restricted/sensitive data will be regulated by the SC and subject to the file creator evaluation. In case patent is granted patent needs to be respected.
	C2	How will the data be made accessible?	by providing the design/manufacturing drawings to the drill rod manufacturers or Investors
	C3	What methods or software tools are needed to access the data?	a "STEP-file" reader from a 3D software program
	C4	Is documentation about the software needed to access the data included?	No
	C5	Is it possible to include the relevant software (e.g. in open source code)?	No
	C6	If there are restrictions on use, how will access be provided?	a strategy concerning also the creation of an Agreement with the licensee to be signed in advance will be considered in WP5.
	C7	How will the identity of the person accessing the data be ascertained?	It will be defined once a strategy related to the licensee agreement will be discussed.

TOPIC	Q. No		
Making data interoperable	D1	Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?	Yes, all data are interoperable. It is very easy to exchange these datas by "STEP-files". Nearly every 3D-software programm can read "STEP-files".
	D2	What data and metadata vocabularies, standards and methodologies will you follow to make your data interoperable?	"STEP-files", a standard in the world of 3D design including architecture, manufacturing, and printing, are used, so no interoperability issues are expected.
	D3	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?	No unavoidable or uncommon drawings are used. If, in case, this is found to be applicable to the datasets generated in WP1, we will provided vocabularies and ontologies to facilitate data mapping.

TOPIC	Q. No		
Increase data re-use (through clarifying licences)	E1	How will the data be licensed to permit the widest re-use possible?	Data classified as ORD can be re-used thanks to open access repository, publications, presentations at seminars, workshops. Data sensible for patents will be subjected to license agreement (to be defined) in order to guarantee the beneficiary IPR and rights to commercialization.
	E2	When will the data made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	The data will be made available for re-use as function of time needed to grant patent
	E3	Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.	After the end of the project, the data can be used by third parties under restrictions. In case patent is granted, the patent needs to be respected.
	E4	How long is intended that the data remains re-usable?	These data can continue to be used by the beneficiary for several years, in order to improve the design, manufacturing and efficiency of the laser drill rod business. The access for third parties will be regulated by the IPR and patents regulations, once they will be defined.

DATASET SHORT NAME	Unique short yet descriptive name of this dataset	Laser drill rods
DATASET REFERENCE IN REPOSITORY (i.e. DOI)	Digital Objective Identifier assigned by Zenodo	Not available at this time
DATASET DESCRIPTION	Long description of this dataset, including detail on variables and parameters, material and methods, time, location and condition of the collecting process (if applicable), reasons that lead to data acquisition actions.	This dataset is generated by Prevent GmbH in order to design and manufacture laser drill rods (a complete laser drill string)
		Not available at this time
STANDARDS AND METADATA	If applicable, standard of measurement and data exchange protocols are described.	Part of the data produced and/or used in the project will be made open in case of scientific publications and to support the results obtained. The data related to the laser drill rods manufacturing are sensible for patent requirements, so they are not accessible at this time. Before defining an access restriction policy (WP5), the drawings need to be prototyped and laboratory testing performance need to be performed.
		The data format are "STEP-files". In order to access the data the following software will be needed: all commercially available 3D drawing programs (3D software); whereby the data exchange format is a "STEP-file".
DATA SHARING	Information regarding shareability and access right of this specific dataset according to Zenodo (i.e. Open Access, Embargoed Access, Restricted Access, Closed Access). Indication of License of use.	Once a patent strategy is defined, access to the laser drill rods data will be defined, specifying what will be openly available and what will be restricted with license restriction. Engineers, researchers and experts in geothermal energy and Investors are the main potential re-users of this dataset.
AFFECTED DOCUMENTS (WPs, Deliverables and OA Publications)	WPs and Open Access Publication that make use of this specific dataset	DeepU Deliverable WP1.

WP2 – FhG-IAPT

TOPIC Q. No

General information	A1	Please, enter your DeepU partner name	FhG-IAPT
	A2	At this time, did your actions during the project generated any research data?	Yes
	A3	At this time, did you used or modified any data acquired from the outside of the project?	Yes
	A4	At this time, did you published any open access scientific peer-reviewed publication that includes research data?	No
	A5	At this time, how many dataset have you generated, used or modified?	5 datasets
	A6	Do you plan to generate even more research data during the lifespan of the project?	Yes
	A7	Do you plan to use or modify data acquired from the outside of the project during the lifespan of the project?	Yes
	A8	Do you plan to publish open access to sceintific peer-reviewed publications that make use of research data during the lifespan of the project?	Yes
	A9	How many dataset do you plan to generate, use or modify during the whole lifespan of the project?	10 datasets

TOPIC Q. No

Data summary	B1	What is the short name that you would use to describe this specific dataset?	Laser drill head and laser and gas process setup
	B2	What is the purpose of the data collection/generation and its relation to the tives of the project?	The data sets are absolutely necessary for the implementation of the laser and gas process for drilling.
	B3	What type and formats of data will the WP generate/collect?	PDF-files and STEP-files (STEP means „Standard for the Exchange of Product model data“ according to ISO Norm 10303) for the design and the drawings of the drill head. Excel-files for the process parameters and other measured values as well as thermographic pictures.
	B4	Will you re-use any existing data and how?	Yes, Fraunhofer IAPT will re-use its own design rules for additive manufactured components and its parameter know-how about rock melting by laser, that have exist before the DeepU project.
	B5	What is the origin of the data?	Fraunhofer IAPT is an expert for additive manufacturing and has a lot of experience in designing 3D printed components with integrated fluid channels as for the drill head is needed.
	B6	What is the expected size of the data?	Design models of the drill head and collected process data < 100 MB
	B7	To whom might it be useful ('data utility')?	The data might be interesting for investors or manufacturer who want to produce a laser drill head. Of course, everyone who wants to repeat the process needs the parameter setup.

TOPIC	Q. No		
Data summary	C1	Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.	Data sensible for patent requirements, as those related to the laser drill head will be subjected by access restriction, respecting the beneficiary's legitimate interests, including those regarding commercial exploitation. A patent strategy will be defined under WP5 and the access to restricted/sensitive data will be regulated by the SC and subject to the file creator evaluation. In case patent is granted patent needs to be respected.
	C2	How will the data be made accessible?	If there are no patent restrictions against, the data will be provided as technical drawings to investors or manufacturers.
	C3	What methods or software tools are needed to access the data?	A "STEP-file" reader of a CAD program is needed.
	C4	Is documentation about the software needed to access the data included?	No
	C5	Is it possible to include the relevant software (e.g. in open source code)?	No
	C6	If there are restrictions on use, how will access be provided?	A strategy concerning also the creation of an agreement with the licensee to be signed in advance will be considered in WP5.
	C7	How will the identity of the person accessing the data be ascertained?	It will be defined once a strategy related to the licensee agreement is discussed.

TOPIC	Q. No		
Making data interoperable	D1	Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?	The data is interoperable. "STEP-files can be exchanged with standard CAD Software. The recorded process data can also be read with standard office software
	D2	What data and metadata vocabularies, standards and methodologies will you follow to make your data interoperable?	Using standard software, Fraunhofer IAPT does not expect any problems exchanging data like STEP, PDF or Excel files.
	D3	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?	No unavoidable or uncommon drawings are used. If, in case, this is found to be applicable to the datasets generated in WP2, Fraunhofer IAPT will provided vocabularies and ontologies to facilitate data mapping.

TOPIC	Q. No		
Increase data re-use (through clarifying licences)	E1	How will the data be licensed to permit the widest re-use possible?	The data classified as ORD can be re-used thanks to open access repository, publications, presentations at seminars and workshops. Data that is sensible for patents will be subjected to license agreement (to be defined) in order to guarantee the beneficiary IPR and rights to commercialization.
	E2	When will the data made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	The data will be made available for re-use as function of time needed to grant patent
	E3	Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.	After the end of the project, the data can be used by third parties under restrictions. In case patent is granted, the patent needs to be respected of course.
	E4	How long is intended that the data remains re-usable?	These data can continue to be used by the beneficiary for several years, in order to improve the design, manufacturing and efficiency of the laser drilling business. The access for third parties will be regulated by the IPR and patents regulations, once they are determined.

DATASET SHORT NAME	Laser drill head and laser and gas process setup
DATASET REFERENCE IN REPOSITORY (i.e. DOI)	Not available at this time
DATASET DESCRIPTION	This dataset is generated by Fraunhofer IAPT in order to design and manufacture the laser drill head for the laser and gas application. Furthermore, the dataset contains the parameter setup for this application.
STANDARDS AND METADATA	<p>Not available at this time</p> <p>Part of the data produced and/or used in the project will be made open in case of a scientific publication, and the data necessary to justify the results obtained will be published.</p> <p>The data related to the design of the laser drill head and its additive manufacturing is sensible for patent requirements, so the data is not accessible at this time. Before defining an access restriction policy (WP5), the drawings need to be prototyped and further laboratory testings need to be performed.</p> <p>The data format are "STEP-files" for the drill head design and "pdf, csv or Excel files" for the parameters.</p> <p>In order to access the design data a CAD software is needed for the "STEP files". Nearly every commercially available 3D drawing program provides the data exchange via "STEP file".</p>
DATA SHARING	<p>Once a patent strategy is defined, access to the laser drill head data will be defined, specifying what will be openly available and what will be protected by license restriction.</p> <p>Engineers, researchers and experts in geothermal energy and investors are the main potential re-users of this dataset.</p>
AFFECTED DOCUMENTS (WPs, Deliverables, and OA Publications)	DeepU Deliverable WP2

WP3 - UNIPD

TOPIC Q. No

General information	Q. No	Question	Answer
	A1	Please, enter your DeepU partner name	UNIPD
	A2	At this time, did your actions during the project generated any research data?	Yes - The WP3 research data collected since the beginning of the WP are starting to be organized in a dataset collecting the whole visual documentation of the samples received so far and in 2 or more datasets collecting preliminary information about the samples (mineralogy, rocks' physical properties, thermal properties, ...)
	A3	At this time, did you used or modified any data acquired from the outside of the project?	No
	A4	At this time, did you published any open access scientific peer-reviewed publication that includes research data?	No
	A5	At this time, how many dataset have you generated, used or modified?	To date two dataset are being created related to the visual documentation of the samples and to the preliminary mineralogical petrographic results
	A6	Do you plan to generate even more research data during the lifespan of the project?	Yes - we plan to generate mineral, petrographical, thermal and mechanical data for fresh, melted only, melted and vitrified rock samples to be organized in different datasets.
	A7	Do you plan to use or modify data acquired from the outside of the project during the lifespan of the project?	Yes - we plan to use as reference some petrophysical data already available, for example, under the P ³ - PetroPhysical Property and the Geo4civhic Project databases
	A8	Do you plan to publish open access to scientific peer-reviewed publications that make use of research data during the lifespan of the project?	Yes - in the following phases of the project, the outputs both of the minero-petrographic and of the thermo-mechanical analyses may be published
	A9	How many datasets do you plan to generate, use or modify during the whole lifespan of the project?	The plan is to generate at least 3 datasets associated each to an unique body of work, as (1) the visual documentation, (2) the minero-petrographic and (3) the thermo-mechanical characterizations of the rock samples

TOPIC Q. No

Data summary	Q. No	Question	Samples' visual documentation	Minero-petrographic data	Thermo-mechanical data
	B1	What is the short name that you would use to describe this specific dataset?	Samples' visual documentation	Minero-petrographic samples' characterization	Thermo-mechanical samples' characterization
	B2	What is the purpose of the data collection/generation and its relation to the objectives of the project?	Sample images collected during different laboratory phases, allow to record the alteration evolution within the samples (unaltered, melted and vitrified zones). In this way the visual information will support the minero-petrographic and thermo-mechanical data interpretation.	The purpose are to define the mechanisms of the rocks /laser/cryogenic gas interaction, to describe the alteration / integrity of the materials, to assess the rocks vitrification and vaporization in laboratory samples	The purpose is to assess the differences between the thermal and mechanical properties of the rock samples before and after the interaction with the laser and the laser + cryogenic gas.
	B3	What type and formats of data will the WP generate/collect?	image data (JPEG, PNG, GIF, TIFF)	Excel - spreadsheet and/ or Word/PDF document - Text	Excel - spreadsheet and/ or Word/PDF document - Text
	B4	Will you re-use any existing data and how?	Previous studies on melted rocks will be considered for data analysis and interpretation.	Reference from existing database will be considered (i.e. P ³ - PetroPhysical Property and Geo4civhic Project databases)	Reference from existing database will be considered (i.e. P ³ - PetroPhysical Property and Geo4civhic Project databases)
	B5	What is the origin of the data?	Mainly internal project data from WP2 and WP3	Mainly internal project data from WP2 and WP3	Mainly internal project data from WP2 and WP3
	B6	What is the expected size of the data?	<10MB	<10MB	<10MB
	B7	To whom might it be useful ('data utility')?	Project partners, future stakeholders, researchers, geothermal and oil and gas companies	Project partners, future stakeholders, researchers, geothermal and oil and gas companies	Project partners, future stakeholders, researchers, geothermal and oil and gas companies

TOPIC	Q. No		
Data summary	C1	Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.	Open publication of data from the 3 datasets is encouraged during the project lifetime. The research data will be made available after the main findings from the final approved research dataset will be accepted for publication and/or 6 months after the laboratory measurements, are completed, according to the project workplan. If some data will be sensible for patent requirements, access restriction will be applied.
	C2	How will the data be made accessible?	Available data will be made accessible through the DeepU project website, the chosen repository (Zenodo) and Open Access publications. Restricted access is foreseen for data containing potential for patentability (access under request evaluated by the SC through the VRE website project section or the Zenodo platform).
	C3	What methods or software tools are needed to access the data?	The data will be provided as standard file (i.e. text, csv). Therefore, to access or read the data open source code or common commercial software (i.e. Microsoft Office) can be used.
	C4	Is documentation about the software needed to access the data included?	None
	C5	Is it possible to include the relevant software (e.g. in open source code)?	Not applicable
	C6	If there are restrictions on use, how will access be provided?	Restricted data access will be provided through the Zenodo digital platform or the VRE project website section.
	C7	How will the identity of the person accessing the data be ascertained?	The access to restricted/sensitive data will be regulated by the SC and subject by the file creator approval through a direct request to the Consortium. The identity of the project partner accessing the data will be ascertained via website login through the private section.

TOPIC	Q. No		
Making data interoperable	D1	Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?	The data produced in WP3 adhere to standards for formats and therefore are expected to be compliant with available (open) software applications, facilitating the re-combinations with different datasets from different origins (i.e. all data included in CSV files will be interoperable.)
	D2	What data and metadata vocabularies, standards and methodologies will you follow to make your data interoperable?	Standard format will be used (i.e. PDF, CSV, DOC, PNG for standard format, CC0 for metadata ...).
	D3	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?	Yes - if this is found to be applicable to the datasets generated in WP3, we will provide vocabularies and ontologies to facilitate data mapping

TOPIC	Q. No		
Increase data re-use (through clarifying licences)	E1	How will the data be licensed to permit the widest re-use possible?	Under Open Access conditions, the latest version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC 0) Licence are considered.
	E2	When will the data made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	The data will be made available at the time of publication or presentation. If an embargo status is defined, an end date for the embargo will be approved by the SC.
	E3	Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.	Yes
	E4	How long is intended that the data remains re-usable?	Indefinite

DELIVERABLE D7.2

Detailed Project Management Plan

DATASET SHORT NAME DATASET REFERENCE IN REPOSITORY (i.e. DOI)	Unique short yet descriptive name of this dataset	Samples' visual documentation	Minero-petrographic data	Thermo-mechanical data
DATASET DESCRIPTION	Digital Objective Identifier assigned by Zenodo	Not available at this time	Not available at this time	Not available at this time
	Long dexcription of this dataset, including detail on variables and parameters, material and methods, time, location and condition of the colleting process (if applicable), reasons that lead to data acquisition actions.	UNIPD team, CNR team	UNIPD team, CNR team	UNIPD team, CNR team
STANDARDS AND METADATA	If applicable, standard of measurement and data exchange protocols are described.	Referring and identification of regulatory conditions applicable to the project technologies	Project technology process analysis, identification of potential failures and identification of mitigation measures to improve process outcome	EHS assessment of laser and cryogenic drilling solutions. Definition of mitigating measures to ensure EHS compliance with applicable regulations, standards operational Health and Safety
		Dataset not available at this time (not completed)	Dataset not available at this time (not completed)	Dataset not available at this time (not completed)
		The data produced and/or used in the project will be made publicly available.	The data produced and/or used in the project will be made publicly available.	The data produced and/or used in the project will be made publicly available.
		The data will be made accessible via Zenodo digital platform, Open Access publication and project website platform.	The data will be made accessible via Zenodo digital platform, Open Access publication and project website platform.	The data will be made accessible via Zenodo digital platform, Open Access publication and project website platform.
DATA SHARING	Information regarding shareability and access right of this specific dataset according to Zenodo (i.e. Open Access, Embargoed Access, Restricted Access, Closed Access). Indication of License of use.	The data format are image data (JPEG, PNG, GIF, TIFF) and/or Word/PDF document.	The data format is Excel - spreadsheet, CSV, and/or Word/PDF document - Text.	The data format is Excel - spreadsheet, CSV, and/or Word/PDF document - Text.
		Microsoft Office / Open Office	Microsoft Office / Open Office	Microsoft Office / Open Office
		No restriction on use is expected.	No restriction on use is expected.	No restriction on use is expected.
		The identity of the person accessing the data will be ascertained via website login through the private section / others / under Zenodo protocol ...	The identity of the person accessing the data will be ascertained via website login through the private section / others / under Zenodo protocol ...	The identity of the person accessing the data will be ascertained via website login through the private section / others / under Zenodo protocol ...
		In order to permit the widest re-use possible the output data are expected to be public.	In order to permit the widest re-use possible the output data are expected to be public.	In order to permit the widest re-use possible the output data are expected to be public.
AFFECTED DOCUMENTS (WPs, Deliverables and OA Publications)	WPs and Open Access Publication that make use of this specific dataset	The data could be used by third parties, in particular after their publication in OA journals and at the end of the project.	The data could be used by third parties, in particular after their publication in OA journals and at the end of the project.	The data could be used by third parties, in particular after their publication in OA journals and at the end of the project.
		The data will remain re-usable at undetermined time.	The data will remain re-usable at undetermined time.	The data will remain re-usable at undetermined time.
		Engineers, researchers and experts in geothermal energy are the main potential re-users of this dataset.	Engineers, researchers and experts in geothermal energy are the main potential re-users of this dataset.	Engineers, researchers and experts in geothermal energy are the main potential re-users of this dataset.
		WP3, D3.1, D3.2, D3.3	WP3, D3.1, D3.2, D3.3	WP3, D3.1, D3.2, D3.3

WP4 – GEOSERV

TOPIC Q. No

General information	A1	Please, enter your DeepU partner name	TERRA GEOSERV LTD (GEOSERV)
	A2	At this time, did your actions during the project generated any research data?	Yes - a review of current legislative and regulatory procedures governing deep drilling and the development of deep geothermal projects
	A3	At this time, did you used or modified any data acquired from the outside of the project?	Yes - Publicly available legislative and regulatory documents for the control and development of geothermal systems and deep drilling operations in Italy and Ireland have been reviewed and accessed. Public deliverables for the Geo-Envi and Optidrill projects are also referenced in the review work
	A4	At this time, did you published any open access scientific peer-reviewed publication that includes research data?	No
	A5	At this time, how many dataset have you generated, used or modified?	1 dataset - a report on the review of existing legislation, regulation and standards is being generated
	A6	Do you plan to generate even more research data during the lifespan of the project?	Yes - we plan to develop a Failure Mode and Effect Analysis (FMEAs), a live EHS risk register for the DeepU technologies and a recommendations document
	A7	Do you plan to use or modify data acquired from the outside of the project during the lifespan of the project?	No - no direct modification of existing data outside the project
	A8	Do you plan to publish open access to scientific peer-reviewed publications that make use of research data during the lifespan of the project?	Yes - in the later phases of the project, the outputs of the FMEA and EHS recommendations may be published
	A9	How many dataset do you plan to generate, use or modify during the whole lifespan of the project?	4 main datasets will be generated in line with the project tasks and deliverables

TOPIC Q. No		Legislative & Regulatory Analysis	Failure Mode and Effect Analysis (FMEA)	Risk register & Assessment Matrix	EHS & Regulatory Recommendations	
Data summary	B1	What is the short name that you would use to describe this specific dataset?	Legislative & Regulatory Analysis Report	Failure Mode and Effect Analysis (FMEA)	EHS Risk Assessment	EHS & Regulatory Recommendations Report
	B2	What is the purpose of the data collection/generation and its relation to the lives of the project?	Referring and identification of regulatory conditions applicable to the project technologies	Project technology process analysis, identification of potential failures and identification of mitigation measures to improve process outcome	EHS assessment of laser and cryogenic drilling solutions. Definition of mitigating measures to ensure EHS compliance with applicable regulations, standards operational Health and Safety	Technology and process recommendations to integrate laser cryogenic drilling into the geothermal and deep drilling sector
	B3	What type and formats of data will the WP generate/collect?	Word/PDF document - Text	Excel - spreadsheet	Excel - spreadsheet	Word/PDF document - Text
	B4	Will you re-use any existing data and how?	Internal data generated by the project on the development of the laser technology and cryogenic gas usage will be referenced in the context of published regulations and standards	Internal Laser and Cryogenic drill string design and operational data (WP1 and WP2). Information from wp3 on rock interaction and vitrification	Internal Laser and Cryogenic drill string design and operational data (WP1 and WP2). Information from wp3 on rock interaction and vitrification	Outputs of WP1, WP2 and WP4 to develop recommendations on technology operational options
	B5	What is the origin of the data?	Government publications. National and International standard bodies	Internal project deliverables from WP1 and WP2	Internal project deliverables from WP1, WP2 and WP3	Internal project deliverables from WP4
	B6	What is the expected size of the data?	<10MB	<10MB	<10MB	<10MB
	B7	To whom might it be useful ('data utility')?	Project partners,	Internal project partners	Internal Project partners	Internal Project partners

TOPIC	Q. No		
Data summary	C1	Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.	The legislative and regulatory analysis relating to deep drilling and geothermal systems (D4.1) will be open data (ORD). The FMEA and EHS risk assessment developed as part of T4.2 (D4.2) as well as the EHS Technology recommendations will be restricted as a result of the patentable technologies developed by the project
	C2	How will the data be made accessible?	Publicly available Data (D4.2) will be made available through the project website. Restricted data will be made available through the Zenodo VRE project platform following the project naming and metadata conventions agreed in the DMP.
	C3	What methods or software tools are needed to access the data?	Microsoft Office
	C4	Is documentation about the software needed to access the data included?	N/A
	C5	Is it possible to include the relevant software (e.g. in open source code)?	N/A
	C6	If there are restrictions on use, how will access be provided?	Restricted data access will be provided through the Zenodo VRE for the project to the project consortium
	C7	How will the identity of the person accessing the data be ascertained?	The access to restricted/sensitive data will be regulated by the SC and subject by the creator of the file through a direct request to the Consortium. The identity of the project partner accessing the data will be ascertained via website login through the private section.

TOPIC	Q. No		
Making data interoperable	D1	Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?	Data from the legislative and regulatory analysis (D4.1) will be classified as ORD and stored in an open access repository and identifiable with a unique reference code (DOI). The FMEA and EHS risk assessment (D4.2) will be RD not classifiable as ORD due to commercial exploitability reasons. This will include the EHS recommendations report (D4.3)
	D2	What data and metadata vocabularies, standards and methodologies will you follow to make your data interoperable?	Standard format will be used (i.e. PDF, CSV, DOC, JPEG ...).
	D3	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?	Yes - if this is found to be applicable to the datasets generated in WP4, we will provide vocabularies and ontologies to facilitate data mapping

TOPIC	Q. No		
Increase data re-use (through clarifying licences)	E1	How will you the data be licensed to permit the widest re-use possible?	Data from the legislative an regulatory analysis (D4.1) classified as ORD can be re-used thanks to open access repository, publications, presentations at seminars, workshops.
	E2	When will the data made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made avaiabel as soon as possible.	At the time of publication or presentation
	E3	Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.	Yes
	E4	How long is intended that the data remains re-usable?	Indefinite

	Unique short yet descriptive name of this dataset	Legislative & Regulatory Analysis	Failure Mode and Effect Analysis (FMEA)	Risk register & Assessment Matrix	EHS & Regulatory Recommendations
DATASET SHORT NAME DATASET REFERENCE IN REPOSITORY (i.e. DOI)	Digital Objective Identifier assigned by Zenodo	Not available at this time R. Pasquai / K. Mallin	Not available at this time R. Pasquai / K. Mallin	Not available at this time R. Pasquai / K. Mallin	Not available at this time R. Pasquai / K. Mallin
DATASET DESCRIPTION	Long description of this dataset, including detail on variables and parameters, material and methods, time, location and condition of the colleting process (if applicable), reasons that lead to data acquisition actions.	Referring and identification of regulatory conditions applicable to the project technologies	Project technology process analysis, identification of potential failures and identification of mitigation measures to improve process outcome	EHS assessment of laser and cryogenic drilling solutions. Definition of mitigating measures to ensure EHS compliance with applicable regulations, standards operational Health and Safety	Technology and process recommendations to integrate laser cryogenic drilling into the geothermal and deep drilling sector
STANDARDS AND METADATA	If applicable, standard of measurement and data exchange protocols are described.	Not available at this time The data produced and/or used in the project will be made publicly available The data will be made accessible via Project VRE - Zenodo Platform The data format is Word/PDF document - Text In order to access the data the following software is needed: Microsoft Office / Open Office	Not available at this time Data will not be made publicly available to protect IP and patents Data will not be accessible to protect IP and patents The data format is Excel - spreadsheet In order to access the data the following software is needed: Microsoft Office / Open Office	Not available at this time Data will not be made publicly available to protect IP and patents Data will not be accessible to protect IP and patents The data format is Excel - spreadsheet In order to access the data the following software is needed: Microsoft Office / Open Office	Not available at this time Data will not be made publicly available to protect IP and patents Data will not be accessible to protect IP and patents The data format is Word/PDF document - Text In order to access the data the following software is needed: Microsoft Office / Open Office
DATA SHARING	Information regarding shareability and access right of this specific dataset according to Zenodo (i.e. Open Access, Embargoed Access, Restricted Access, Closed Access). Indication of License of use.	No restriction on use is expected The identity of the person accessing the data will be ascertained via website login through the private section / others / under Zenodo protocol ... In order to permit the widest re-use possible the output data are expected to be public. The data could be used by third parties, in particular after the end of the project, but some data could be restricted. The data will remain re-usable at undetermined time. Engineers, researchers and experts in geothermal energy are the main potential re- users of this dataset.	Restriction on use is expected	Restriction on use is expected	Restriction on use is expected
AFFECTED DOCUMENTS (WPs, Deliverables and OA Publications)	WPs and Open Access Publication that make use of this specific dataset	WP4, D4.1	WP4, D4.2	WP4, D4.2	WP4, D4.3

WP5 – RED

TOPIC	Q. No		
General information	A1	Please, enter your DeepU partner name	RED
	A2	At this time, did your actions during the project generated any research data?	NO
	A3	At this time, did you used or modified any data acquired from the outside of the project?	NO
	A4	At this time, did you published any open access scientific peer-reviewed publication that includes research data?	NO
	A5	At this time, how many dataset have you generated, used or modified?	NONE
	A6	Do you plan to generate even more research data during the lifespan of the project?	YES
	A7	Do you plan to use or modify data acquired from the outside of the project during the lifespan of the project?	YES
	A8	Do you plan to publish open access to scientific peer-reviewed publications that make use of research data during the lifespan of the project?	NO
	A9	How many dataset do you plan to generate, use or modify during the whole lifespan of the project?	FOUR

TOPIC	Q. No		DATA SET 1: drilling costs	DATA SET 2: levelised electrical and heat cost	DATA SET 3: IPR	DATA SET 4: Numerical modelling
Data summary	B1	What is the short name that you would use to describe this specific dataset?	DEEP DRILLING COSTS: State of art versus Deep-U at different depths	Steam and hot water based power plants: capital and operating costs for 5 and 10 Mwe plant	IPR	Numerical modelling results
	B2	What is the purpose of the data collection/generation and its relation to the tives of the project?	Indicative cost reduction potential of deep-U drilling	Cost of renewable electricity and heat extraction potential from deep U through modelling	Review actual state of art and IPR potential of deep U technologies	Model simulation of underground closed loop heat extraction
	B3	What type and formats of data will the WP generate/collect?	Drilling cost estimates per cost type	Levelised cost prediction per kWh of electricity and heat	Filed patents and patentability of deep U technologies	Modelling output, in form of graphs and textual description (pdf file) and raw modelling output data
	B4	Will you re-use any existing data and how?	Update of existing drilling cost data from literature and drilling operators	Use previous studies of power plants, including Pprevent studies; modeling of deep U energy extraction	Exisuing patent reviews including already filed patents of deep U partners	No
	B5	What is the origin of the data?	Literature and deep drilling operators	Literature and power plant suppliers	Patent literature	Modelling performed in the frame of DeepU
	B6	What is the expected size of the data?	Do not know for the moment	Do not know for the moment	Do not know for the moment	5 Mb
	B7	To whom might it be useful ('data utility')?	Future stakeholders, i.e. drilling contractors and multi-utility companies, oil and gas companies, ventures capitalists, public authorities	Future stakeholders, i.e. multi-utility companies, public authorities	To the consortium partners to define IPR strategy and protect deep U technology	Project partners and geothermal sector for defining economics of the proposed solution

DELIVERABLE D7.2 Detailed Project Management Plan

TOPIC	Q. No		DATA SET 1: drilling costs	DATA SET 2: levelised electrical and heat cost	DATA SET 3: IPR	DATA SET 4: Numerical modelling
Data summary	C1	Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.	Drilling costs state of art and deep U	Levelised electrical and heat cost	None until patent strategy is defiend and eventual patent request are publicly available	All data will be open
	C2	How will the data be made accessible?	Publications, presentations in seminars, stakeholder workgroups	Publications, presentations in seminars, stakeholder workgroups	None until patent strategy is defiend and eventual patent request are publicly available	Uploaded in the chosen repository (ZENODO), presented in public events
	C3	What methods or software tools are needed to access the data?	None	None	None until patent strategy is defiend and eventual patent request are publicly available	Most data can be accessed with open source software. Raw modelling data may (to be defined) require COMSOL Multiphysics software
	C4	Is documentation about the software needed to access the data included?	None	None	None until patent strategy is defiend and eventual patent request are publicly available	No
	C5	Is it possible to include the relevant software (e.g. in open surce code)?	Not applicable	Not applicable	None until patent strategy is defiend and eventual patent request are publicly available	No
	C6	If there are restrictions on use, how will access be provided?	Access to summary data, no detailed background data	Access to summary data, no detailed background data	None until patent strategy is defiend and eventual patent request are publicly available	NA
	C7	How will the identity of the person accessing the data be ascertained?	The access to restricted/sensitive data will be regulated by the SC and subject by the creator of the file through a direct request to the Consortium. The identity of the project partner accessing the data will be ascertained via website login through the private section. The acces to open data (Dataset 4) do not need to ascertain identities since data will be public			

TOPIC	Q. No		DATA SET 1: drilling costs	DATA SET 2: levelised electrical and heat cost	DATA SET 3: IPR	DATA SET 4: Numerical modelling
Making data interoperable	D1	Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?	NO	NO	NO	All data include in CSV files will be interoperable. Numerical simulation output data description will include input data and parameters, to allow result checking by reproducing the same modelling conditions
	D2	What data and metadata vocabularies, standards and methodologies will you follow to make your data interoperable?	NONE	NONE	NONE	CC0 for metadata, standard format will be used (CSV, PDF, DOC, PNG)
	D3	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?	NO	NO	NO	Yes, in case of need

DELIVERABLE D7.2

Detailed Project Management Plan

TOPIC	Q. No		DATA SET 1: drilling costs	DATA SET 2: levelised electrical and heat cost	DATA SET 3: IPR	DATA SET 4: Numerical modelling
Increase data re-use (through clarifying licences)	E1	How will you the data be licensed to permit the widest re-use possible?	Publications, presentations at seminars, workshops	Publications, presentations at seminars, workshops	No re-use until patent requests are publicly available	Open data license (CC BY, CC0)
	E2	When will the data made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	At the time of publication or presentation	At the time of publication or presentation	Function of time needed to grant patent	At the time of publication or presentation
	E3	Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.	Yes	Yes	In case patent is granted patent needs to be respected	Yes
	E4	How long is intended that the data remains re-usable?	Indefinite	Indefinite	In case patent is granted patent needs to be respected	Indefinite

	Unique short yet descriptive name of this dataset	DATA SET 1: drilling costs	DATA SET 2: levelised electrical and heat cost	DATA SET 3: IPR	DATA SET 4: Numerical modelling
DATASET SHORT NAME					
DATASET REFERENCE IN REPOSITORY (i.e. DOI)	Digital Objective Identifier assigned by Zenodo	Not available	Not available	Not available	Not available at this time
DATASET DESCRIPTION	Long description of this dataset, including detail on variables and parameters, material and methods, time, location and condition of the collecting process (if applicable), reasons that lead to data acquisition actions.	Not applicable	Not applicable	Not applicable	Not available at this time
STANDARDS AND METADATA	If applicable, standard of measurement and data exchange protocols are described.	Not applicable	Not applicable	Not applicable	Not available at this time
DATA SHARING	Information regarding shareability and access right of this specific dataset according to Zenodo (i.e. Open Access, Embargoed Access, Restricted Access, Closed Access). Indication of License of use.	Not available	Not available	Not available	No restriction of use is expected
AFFECTED DOCUMENTS (WPs, Deliverables and OA Publications)	WPs and Open Access Publication that make use of this specific dataset	WP5; D5.2	WP5; D5.1 and D5.2	WP5; D5.3	WP3, WP5

Appendix B

GUIDELINES TO UPLOAD DATASETS ON THE ZENODO REPOSITORY

To upload resources (i.e., datasets or pre-proof manuscripts) on the [Zenodo](https://zenodo.org) repository, please follow these instructions to guarantee an accurate description by metadata.

1. General Instructions

In Zenodo, all the fields with the **red asterisk** are **mandatory**. For DeepU contributions, please pay special attention to

- Choose the “Deep Geothermal” Community (point 6 below)
- Choose proper keywords (point 15 and section 2 below)
- Mention the Grant (point 19)

Step-by-Step:

1. Go to the URL: <http://zenodo.org>
2. Login with your own Zenodo **username** and **password** (or register and then login in case you are not yet registered)
3. Click on ‘**Upload**’ (on the top menu near the search text box)
4. Click on ‘**New upload**’ (green button on the top right)
5. In the ‘Files’ frame, drag and drop the pdf of the resource or click on ‘**Choose file**’ to browse your computer and select the pdf file to upload. Then click on ‘**Start upload**’ on the top right of the ‘Files’ frame. When the upload is complete, you will see the size of the file and the “Progress” will be signalled with a
6. **The following is the most crucial step:** In the ‘Communities’ frame, start writing ‘Geothermal’ till the ‘**Deep Geothermal**’ choice appears. Choose ‘Deep Geothermal’
7. In ‘Upload type’, choose the correct typology referred to the uploaded file (“Publication” by default), then define the ‘**Publication type**’ (e.g., Project deliverable, Dataset, ...)
8. If the document already has a **DOI**, insert the existing DOI code in the ‘Digital Object Identifier’; otherwise, leave it blank. In this latter case, Zenodo will assign a new DOI to the document.
9. In ‘**Publication date**’, insert the date of the document delivery (publication)
10. In ‘**Title**’, insert the document title
11. In ‘**Authors**’, insert the family name *comma* given name, the affiliation and, if known, the ORCID code. In case of more than one author, click on ‘+ Add another author’ and fill in the three fields. If the affiliation is not clear, leave the field blank. In case the author/s is/ are not declared in the document, use the name of the Project (i.e., DeepU)
12. In ‘**Description**’, you should describe the resource (free text). The abstract of the pre-proof manuscript should be used. If this is unavailable, please describe the document in a few lines (e.g., for datasets). Pay attention when you copy+paste a text since some lines might wrap and should be corrected. The text has to be correctly formatted
13. In ‘**Version**’, insert the code of the resource version. If the code is not available, use the value 1.

14. **'Language'** has to be English. If you type 'eng', the complete **'English'** word appears and can be chosen.
15. The **'Keyword'** field is essential: Please follow carefully the dedicated instruction provided in Section 2.
16. Leave the **'Additional note'** field empty.
17. In **'Access right'** choose 'Open Access' if documents/datasets are freely accessible, 'Embargoed Access' if documents/datasets will be available by a specific date, 'Restricted Access' if documents/datasets are available under specific conditions or 'Closed Access' if the documents/datasets are not accessible. Consider that in all 4 cases, the metadata will be available for the documents/datasets.
18. If the **'Licence'** is specified in the document/datasets, it should be chosen in the list that appears when you start to type in the dedicated text box. If the document's licence is not declared, type 'not' and 'Licence Not Specified' will appear and have to be clicked. In this latter case, for the DeepU documents/datasets it is warmly suggested to assign a Creative Commons (CC) licence.
19. The next metadata frame is **'Funding'**. The values to be chosen are **'European Commission (EU)'** in the **'Grants'** left field and type the name of the project **'DeepU'** in the right field. If one item has been already inserted for DeepU, you will find it in the list of projects already available and you can choose it. Please check that the funding is properly stated as "European Union's HORIZON-EIC-2021-PATHFINDEROPEN-01 programme under grant agreement No 101096437".
20. Insert **'Contributors'** as described at point 11, also selecting in the right field the contributor's role (e.g., contact person, data collector, data curator, editor, ...)
21. If existing, please insert the **'References'** available in the resource by copy+paste in the specific field, one reference per field and using 'Add another reference' till the last item
22. In case the resource refers to a 'Journal', 'Conference', 'Book/Report/Chapter', 'Thesis', or 'Subject', please complete the description in the proper field
23. **'Save'** your metadata by clicking on the button at the bottom right of the page and checking your work. When you are sure your description is correct, click on **'Publish'**. During the description of the resource, it is recommended to save often, to avoid loss of information

2. Specific instruction to insert Keywords (point 15)

It is mandatory to insert

- all the required keywords as described below
- only one keyword per keyword field, using 'Add another keyword' to insert new keywords
- each keyword in lowercase and without special characters (e.g., -, ...) among the words

Insert:

1. geothermal energy
2. at least one of the following categories¹:
 - a. exploration
 - b. drilling
 - c. production
 - d. surface systems / generation

¹ These categories correspond to the ETIP-DG Working Group (WG) titles

- e. non-technical
 - f. environmental
3. if the category is 'exploration', choose **at least one** topic among: geochemistry, geophysics, structural geology, hydrogeology, geochronology, volcanology (multiple topics are allowed)
 4. if the category is 'non-technical', choose **at least one** topic among: social aspects, codes & definitions, energy sector status, skills education & training, research roadmaps, regulatory aspects, economics, financing (multiple topics are allowed)
 5. name of the location if the document is referred to a particular place
 6. country if the document is referred to a specific nation