D5.3 – Data Management Plan

WP5 – Task 5.2

Version n°0.1

December/2024

Yann PERIN[1], Daria BARSUK[2]

[1] - GRS, [2] - LGI

Disclaimer

Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

Document information

Grant Agreement	n°101163288
Project Title	Experiments for Validation and Enhancement of the REactor preSsure vessel fluence assessmenT
Project Acronym	EVEREST
Project Coordinator	Dr. Yann Perin, GESELLSCHAFT FUR ANLAGEN UND REAKTORSICHERHEIT (GRS) gGmbH
Project Duration	1 September 2024 – 31 August 2028 (48 months)
Related Work Package	WP5
Related Task(s)	T5.2
Lead Organisation	GRS
Contributing Partner(s)	LGI
Due Date	M3 (November 30th 2024)
Submission Date	13/12/2024
Dissemination level	Public

History

Date	Version	Submitted by	Reviewer 1	Reviewer 2	Comments
25/11/24	0.1	Daria Barsuk (LGI)			First draft of the document created
26/11/24		Daria Barsuk (LGI)	Gabor Szendro (LGI)	Yann Perin (GRS)	Minor corrections in the wording and formatting across the text. Specific partners provision, national repositories, key words edit.
04/12/24	0.2	Daria Barsuk (LGI)	All partners		Final Version 0.1 created

Signatures

Author e-Signature	WPL e-Signature	Coordinator e-Signature
Daria BARSUK (LGI) 	Yann PERIN (GRS) 	Yann PERIN (GRS, COO)

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

Table of contents

List of figures

Abbreviations and acronyms

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

Acronym	Description
EEAB	External Expert Advisory Board
CA	Consortium Agreement
COO/co-COO	Coordinator/co-Coordinator
D or DEL	Deliverable
DoA	Description of Action
EC	European Commission
ExCom	Executive Committee
EUG	End Users Group
GA	Grant Agreement
GB	Governing Board
HE	Horizon Europe
KOM	Kick-Off Meeting
IMR	Interim Management Report
M	Month (of the project)
MS	Milestones
NDA	Non-Disclosure Agreement
PMO	Project Management Office
PM	Person-Month
PO	Project Officer
PQP	Project Quality Plan
PR	Periodic Report
RP	Reporting Period
V	Version
WP	Work Package
WPL	Work Package Leader
Gold Open Access	Gold Open access publishing means that an article is immediately provided in open access mode on the publisher or journal's website. Some publishers charge Article Processing Chargers (APCs) to make articles open
Green Open Access	Self-archiving (green open access) means that a published article or the final peer-reviewed manuscript is archived (deposited) in an online repository before, alongside or after its publication. In some cases, the author can choose to delay access to the article (embargo period). Embargo period rules originating from H2020 program state that embargo periods cannot exceed six months, except for publications in social science and humanities where the maximum embargo period is twelve months
Personal Data	Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

	collected together can lead to the identification of a particular person, also constitute personal data. Personal data that has been de-identified, encrypted or pseudonymised but can be used to re-identify a person remains personal data and falls within the scope of the law. Personal data that has been rendered anonymous in such a way that the individual is not or no longer identifiable is no longer considered personal data. For data to be truly anonymised, the anonymisation must be irreversible
Zenodo	Zenodo is a catch-all research data repository that enables researchers, scientists, EU projects and institutions to share research results, make research results citable, and search and reuse open research results from other projects. Zenodo is harvested by the OpenAIRE portal and hosted by the CERN cloud infrastructure

Table 1. Abbreviations and Acronyms

Summary

This deliverable presents the first version of the EVEREST Data Management Plan (DMP). It describes the guiding principle for data management and gives an overview of what data will be gathered and processed in the project, according to the EC FAIR Data Management principle making data findable, accessible, interoperable and reusable.
The purpose of the DMP is to contribute to good data handling through indicating what research data the project expects to generate and describe which parts of the data that can be shared with the public. Furthermore, it gives instructions on naming conventions, metadata structure, storing of the research data and how to make public data available.
This deliverable will be regularly updated during the project life.

Keywords

Data Management, metadata, open data, research data, data repository, data storage, Zenodo, FAIR data, findable, accessible, interoperable, reusable, data format, data storage, GDPR, sensitive data.

1 Introduction

This deliverable provides an easy overview of research data the project is expected to generate, the types and formats of this data, and how this data is processed and stored to make them findable, accessible, interoperable and reusable, according to the principles of FAIR data management. The purpose of the DMP is to contribute to good data handling during the project's lifetime, and to describe how such data will be curated and preserved.

1.1 Objectives and scope of the document

The DMP describes the data management life cycle for the data to be collected, processed and/or generated by EVEREST project, as a Euratom/Horizon Europe project. The DMP aims at defining the management strategy of data generated during the project with the purpose of making research data findable, accessible, interoperable and reusable (FAIR). The DMP addresses the following points:

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

- the handling of research data during and after the end of the project,
- what data will be collected, processed and/or generated,
- which methodology and standards will be applied,
- whether data will be shared/made open access, and
- how data will be curated and preserved (including after the end of the project).

1.2 Versioning

According to the EU's guidelines regarding the DMP, the document may be updated - if appropriate - during the project lifetime (in the form of deliverables). The minimum requirement is that the DMP be updated for each periodic evaluation of the project.
The DMP is intended to be a living document in which information can be made available progressively as the project advances.
DMPs should, therefore, have a clear version number and include a timetable for updates.

2 Data summary

The EVEREST project main objectives are directly linked to increasing the safety of Pressurized Water Reactors with the help of High-Fidelity Multiphysics modelling of the reactor core, particularly to determine the neutron radiation exposure (fluence) affecting the Reactor Pressure Vessel (RPV) . Those principle objectives of this project, for which the main portion of data is collected (see category "Research and Scientific Data" in Table 2), generated or re-used are the following:
1. To provide guidance for the determination of the source term for Reactor Pressure Vessel (RPV) and core internals fluence evaluation in PWR and VVER and to contribute to establish the best practices for the safety analysis in plant life extension scenarios (WP1);
2. To produce validated advanced multi-physics models of the research reactors involved in the project (3 in total: at BME, JSI and at HUN-REN EK) (WP2);
3. To successfully perform multi-physics experiments of benchmark quality for validation of advanced multi-physics models, targeting a better spatial resolution (WP3);
4. Provide training, learning material and educational environment for professionals on multi-physics methods and for nuclear regulators on the advantages of high-fidelity nuclear reactor modelling (WP4);
Meeting these objectives requires cooperative work of all 50 researchers from 15 partnering institutions. Each partner will collect or generate specific data, related to their tasks as per the GA. As a result of preliminary exchanges within the EVEREST consortium, it has become apparent to us that it is challenging to give an exact estimation of the kind and volume of data at this early stage of the project. However, during the course of the Action we are to correct the information given below so that it better reflects the generated research and scientific data, produced by each partner, as well as data produced as a result of other support activities such as project management and dissemination. Nevertheless, some preliminary data summary is provided in the table below.

Table 2 contains responses to the following questions, addressed by all partners:

- What is the purpose of the generated/collected data
- What types and formats of data will the project generate/collect
- Whether any existing data will be re-used and what will be re-used
- What is the origin/provenance of the data
- What is the expected size of the data that the project intends to generate or re-use
- What is the origin/provenance of the data
- What is the data utility - to whom might your data be useful outside the project

Funded by the European Union
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

Data category	Data Purpose	Data Origin	Data Types	Data Formats	Existing data to be used or re-used	Estimated Data Size	Data Utility
Research & Sientific data	In WP1, collected data will be used as origin for the generation of core and ex-core models for 1 VVER and 2 PWR, the calculation of cycles for those reactors, the calculation of fluence and dosimetry surveillance, as well as uncertainty analysis for the cycle calculations. In WP2 and 3, colleted data will be used as origin for the generation of models of 3 research reactors, the simulation of transient, as well as uncertainty analysis. The generated data includes power/flux/TH steady-state profiles, fluence/dosis value and/or time-dependent evolutions.	In WP1, the data will originate from the power plant operators taking part in the project, i.e. PEL for the German PWR, AXPO for the Swiss PWR and PAKS NPP for the VVER 440. In WP2 and 3, the data will originate from the institute where the research reactor is located, namely ISI and BME and HUN-REN.	Code inputs/outputs. In-core data: neutron fission source and other in-core parameter profiles. Ex-core data: fluence maps, dosimetry/surveillance data Power/flux/TH profiles and/or time-dependent evolution Multi-physics steady-state/time-dependent data Detailed CAD models of experiments, high fidelity, time dependent vector data of reactor coolant temperatures and velocities, neutron flux measurements in multiple positions Geometric and material description of fuel and control assemblies, reactor internals and reactor pressure vessel. Cycle specific information, e.g. loading pattern and operational in-core measurements.	ASCII VTK/VTR STP HDF5 binary NPY STL PolyMesh/OpenFOAM .docx .xlsx .csv tsv .json	The power plant data already exists but must be put in a format usable for the partners. In one case, JSI Partner backgroud data from Multiphysics experiments performed by Dr. Romain Henry (currently Framatom) during his PhD thesis, obtained between 2014-2017. The data was obtained at the JSI TRIGA facility.	1 < Data size < 13 TB	OECD/NEA Working Party on Scientific Issues and Uncertainty Analysis of Reactor Systems (WPRS) Other technical organizations, such as TSOs, fuel/reactor vendors and reactor operators.
MGMT & Legal data	Some data will be generated or reused to support the management and monitoring of the project	Data (templates) originates from the a) EU templates and guidelinesl b) Earlier company projects and cumulated description of best practices and gained expertise (only relevant to generic, EVERST non-specific data)	Consortium contacts, Monitoring and reporting files, Surveys, MS Office Templates	.xlsx .docx .pdf .pptx Cryptographic data, such as signatures, proof requests, etc	Several procedural documentation templates, including (legal, management files (CA, Deliverables) describing standard practices and legislation will be re-used in part and adapted to this project	Data size < 500 MB	EU/MS governments, policy makers and institutions EU institutions & Regulators General public Projects/Initiatives : EU funded projects
Dissem. & Commun. Data	Data will be generated to support communication and dissemination of the project results and objectives	Photos, progress reports, project's events, stock images	Partners' logos and promotional images/templates. Web digital content Hard copies	.xlsx .docx .pdf .pptx .jpg .tiff .html .json	No data from previous work will be used or re-used for the communication or dissemination purposes ofthis project	Data size < 500 MB	General public Specialized public: Research peers & Students EU institutions & Regulators Other EU-funded projects
Personal Data	The project is gathering personal data such as email addresses through the newsletter or for the management and coordination purposes of the project. *The website does not use cookies. **The project does not organize recruitment campaigns and does not generate or store candidates data	Collaborators contacts Subscribers contacts Attendees contacts (in case of external events)	Email address/Names Partner Entity Role in the project	.xlsx .json	No data from previous work will be used or re-used	Data size < 10 MB	Data will not be used outside the project Collected data is GDPR complient (see Section 5.2)

Table 2 Estimation of the data that will be collected during the EVEREST project by all partners

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

More information and guidelines to better understand how to properly address the points above could be found via the Digital Curation Centre[1] (DCC). The DCC was founded in 2004 to enable adherence of research data to FAIR principles. The information shared in the cited resource can be seen as a guideline for the implementation of a successful Data Management Plan.

3 Data Management Policy: FAIR data

EVEREST's general data management policy that is presented in the subsequent chapters has been developed in accordance with Horizon Europe FAIR principles (Findable, Accessible, Interoperable and Reusable), open data requirements and implementation guidelines. It applies mainly to new results that are produced in EVEREST and that are to be made available by the project consortium as open source, open science and open data.

3.1 Making data findable

Template guiding questions:
Will data be identified by a persistent identifier?
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, outline what type of metadata will be created and how.
Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
Will metadata be offered in such a way that it can be harvested and indexed?

Digital Object Identifier (DOI)
All open data, publications, and open-source software generated within the EVEREST project will be identifiable through persistent Uniform Resource Locators (URLs). Whenever possible, these results will also be assigned a Digital Object Identifier (DOI) to ensure they are easily and uniquely citable.
For open results stored in the default Open Access repository (Zenodo, as described below), a DOI will be automatically generated, along with the added benefit of Zenodo's DOI versioning capabilities.
Open results deposited in institutional repositories, publisher repositories, or other data and research archives will, at a minimum, have persistent URLs for identification. If the hosting institution collaborates with a DOI registration agency, a DOI will also be assigned.
The assignment of unique identifiers such as DOI, Publisher Item Identifier (PII), or International Standard Serial Number (ISSN) to scientific publications will depend on the open access approach (green or gold) selected by the authors. It will also vary based on the policies of the chosen scientific publisher and research repository.
Metadata
In order to keep data findable, it is necessary to provide its metadata. Metadata is a systematic data on the research data themselves. Author, date created, date modified and file size are examples of very basic document metadata.

Metadata enables other researchers to find data in an online repository and is, as such, essential for the reusability of the dataset. By adding rich and detailed metadata, other researchers, can better determine whether the dataset is relevant and useful for their own

[1] https://www.dcc.ac.uk/guidance

Funded by the European Union
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

research. Metadata will be uploaded in a standardized form. This metadata will be kept separate from the original raw research data.

The following deposition metadata fields are mandatory for EVEREST:

- the terms "European Union (EU)" and "Horizon Europe";
- the name of the action, acronym and grant number;
- the title and description of the deposition metadata,
- the upload type (publication, dataset, software, …),
- the publication date (ISO8601 format, YYYY-MM-DD), and length of embargo period if applicable;
- the creators/authors of the deposition,
- the persistent identifier (DOI),
- keywords.

Zenodo

The EVEREST project will store its open data in Zenodo, an open-access online repository designed for research data. Zenodo's structure, features, and management adhere to FAIR data principles. As a service provided by OpenAIRE and hosted by CERN, Zenodo allows researchers to deposit both publications and datasets, offering tools to connect them using persistent identifiers and enabling proper data citation.

Zenodo is specifically designed to support the core principles of FAIR–making data findable, accessible, reusable, and interoperable–which are essential for Open Research Data (ORD) projects. It is a versatile repository that accommodates researchers, scientists, EU-funded initiatives, and institutions by offering the following benefits:

- Facilitating the sharing of research outputs in various formats, including text, spreadsheets, audio, video, and images, across all scientific disciplines.
- Enhancing visibility and recognition of research by making outputs citable and seamlessly integrating them into reporting systems for funding bodies such as the European Commission.
- Enabling easy access to and reuse of shared datasets.
- Automatically generating Digital Object Identifiers (DOIs) for all deposits.
- Connecting research outputs to the OpenAIRE portal for broader accessibility and integration.

Zenodo serves as a comprehensive platform to ensure EVEREST's research data is well-managed, accessible, and aligned with open science principles.

Search keywords

Zenodo allows to perform simple and advanced search queries on Zenodo using the keywords. Zenodo also provides a user guide with easy to understand examples. The Data Controllers at each pilot site will be responsible for uploading public datasets that they have generated and to assign specific keywords relevant to these datasets. Dataset specific keywords must be descriptive to the content of the dataset. In addition, the consortium will define a set of general keywords that should apply to corresponding public datasets, scientific publications and public deliverables. Example of such key words:

- EVEREST, VVER, PWR, High-Fidelity, Multi-physics experimental data, High-Fidelity, Multi-physics validation, Safety, Codes, Reactor, Communication, Dissemination

Other repositories

Examples of other repositories, that will be used to make data openly available include HAL (https://hal.archives-ouvertes.fr) and potentially other national repositories known by the partners.

Naming conventions

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

Files and folders at data repositories will be versioned and structured by using a name convention consisting as follow: FileType_EVEREST_[name of the document]_Vxy_date_. FileExtension.

Version numbers

Individual file names and datasets will contain version numbers that will be incremented at each revision (Vxyz). For publications, versioning is in general not necessary.
Zenodo provides DOI versioning of all datasets uploaded to their communities, which allows us to edit and update the uploaded datasets after they have been published. This also allows us to cite specific versions of an upload and cite all versions of an upload.

3.2 Making data openly accessible

Template guiding questions:
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.
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.
Will the data be accessible through a free and standardized access protocol?
If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
How will the identity of the person accessing the data be ascertained?
Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?
Repository:
Will the data be deposited in a trusted repository?
Have you explored appropriate arrangements with the identified repository where your data will be deposited?
Does the repository ensure that the data is assigned an identifier?
Will the repository resolve the identifier to a digital object?
Metadata:
Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why.
Will metadata contain information to enable the user to access the data?
How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?
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)?

3.2.1 By default openly available data

In order to maximise the impact of EVEREST research data, the results are shared within and beyond the consortium. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through open access data repositories.

Moreover, in EVEREST, we have strived to make as many deliverables public as possible: 18 out of 23 deliverables will be open for public access.

The EVEREST project datasets are first stored and organized in a database by the data owners (personal computer, or on the institutional secure server) and on the project's website. All data are made available for verification and re-use, unless the task leader can justify why data cannot be made openly accessible. To protect the copyright of the project knowledge, Creative Commons license will be used when needed.

The EVEREST public datasets and deliverables (data access policy unrestricted) will be accessible by:

- EVEREST project web site https://www.projecteverest.eu/
- Partners database
- OpenAIRE
- Zenodo and its national homologues for ORDP data and datasets

Funded by the European Union
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

- Open access journals

All data deposited on Zenodo are accessible without restriction for public.

Exceptions are the following:

- Copyright and permissions for re-using third-party data sets
  Processing and combining input data from many different sources may lead to unclear IPR situations regarding the generated output data, therefore such repurposed data (e.g. model output data) can only be made open if any of the underlying data (e.g. model input data) is open, too.
  For this: potential users must contact the IPR (Intellectual Property Rights) team or the data owner in order to gain access. If necessary, appropriate IPR procedure (such as non- disclosure agreement - NDA) will be used.
- Personal data treatment and confidentiality issues
  Datasets referring to the quality and quantity of certain elements at risk, such as people and critical infrastructures, are not open by default as their publication may pose privacy, ethical or security risks.
- Data-driven business model
  Data that will be exploited commercially will not be made open.

3.2.2 Data Accessibility

EVEREST open results will be made accessible according to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon Europe.

Open data: All open results (data, software, public documents, scientific publications) of the project will be openly accessible at an appropriate Open Access repository (i.e. Zenodo) as soon as possible. Specifically, research data needed to validate the results in the scientific publications will be deposited in a data repository at the same time as a publication. Non-public research data will be archived at the repository using a restricted access option.

Scientific publications: Providing open access to peer-reviewed scientific publications can be ensured either by publishing in green or gold open access journals with or without author processing fees. Any scientific publications from EVEREST and the related bibliographic metadata must be made available as open access and announced on the project website as well as in the OpenAIRE portal (https://www.OpenAIRE.eu/) and the R&I Participant Portal (https://ec.europa.eu/research/participants). To automate the process of reporting scientific publications and related research data in OpenAIRE, the publication should be deposited in an OpenAIRE-compliant repository, either by the authors of the publication (green open access) or by a scientific publisher (gold open access). While additional forms of disseminating open access papers, including academic social network sites such as ResearchGate (https://www.researchgate.net/) are possible, an electronic copy of the publication has to be deposited in suitable open access repository in the first place. According to the European Research Council's Guidelines on Open Access, "Venues such as Research Gate or Academia.edu that require users to register in order to access content do not count as repositories. The posting of publications on a personal, institutional or project specific webpage or the deposit in a publicly accessible Dropbox account is not sufficient to satisfy the requirements either."
If the chosen repository is not fully OpenAIRE compliant, the publications or data must be linked at https://www.openaire.eu/participate/claim with the respective funding agency (European Commission in this case). Green open access journals or gold open access journals without author processing fees should be preferred for disseminating scientific

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

publications of the EVEREST project. Nevertheless, the journal's visibility and prestige (translated in the Impact Factor), together with the speed of publication, should be considered when choosing a journal for publication of a manuscript. According to the EC recommendation, authors of the publication are encouraged to retain their copyright and grant adequate licences to publishers.

- Green open access (self-archiving)
  Green open access or self-archiving means that the published article or the final peer-reviewed manuscript is archived by the researcher itself in an online repository, in most cases after its publication in the journal. The journal must grant the researcher the permission to self-archive the final peer-reviewed article, at the latest, 12 months after publication. For finding suitable green open access publishers, researchers are encouraged to consult RoMEO (http://sherpa.ac.uk/romeo), a searchable database of publisher's policies regarding the self- archiving of journal articles on the web and in Open Access repositories.
- Gold open access (open access publishing)
  Gold open access means that the publication is available by the scientific publisher as open access. Some journals require an author-processing fee for publishing open access. Author-publishing fees for gold open access journals can be reimbursed within the project period and budget. Some publishers allow the researcher to deposit a copy of the article in a repository, sometimes with an embargo period. For finding suitable gold open access publishers, researchers are encouraged to consult the Directory of Open Access Journals (https://doaj.org/), a service that indexes high quality, peer-reviewed open access academic journals that use an appropriate quality control system.

Regarding the mere access to open data deposited as data files in a data repository, there are no special methods or software tools needed. The files can be downloaded from the data repository using a standard web browser. The offline viewing, interpreting, processing and editing of data files downloaded from the data repository, it heavily depends on the type and format of the data.

3.2.3 Specific partner provisions

- PEL, AXPO and PAKS: Non-sensitive data produced by those partners within WP 1 will be made available for public during and after the project. Sensitive data will be made available for partners upon request during the project as described in the Consortium Agreement. The data transfer process will comply with the applicable export control laws.

Every partner, who will institute a restricted access to their datasets, will be asked to install a protocol where the identified collaborators will be explained how to access these datasets. If applicable, this will be mentioned in the next version of the Data Management Plan.

3.2.4 How access will be provided if there are restrictions on use to data

Where a restriction on open access to research data is necessary, attempts will be made to make data available under controlled conditions to other individual researchers. In the case where restricted or embargoed data is stored in the Zenodo repository, information about the restricted data will be published in the repository, and details of when the data will become available will be included in the metadata. According to the Q&A session "Open Research Data in HE and Zenodo repository", Metadata for both open, closed, embargoed

Funded by the European Union
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

and restricted records are always publicly available in Zenodo. Data files and data sets for restricted access records are only visible to their owners and to those the owner grants access. Restricted access allows a researcher to upload a dataset and provide the conditions under which he/she grants access to the data. Researchers wishing to request access must provide a justification for how they fulfil these conditions. The owner of the dataset gets notified for each new request and can decide to either accept or reject the request. If the request is accepted, the requester receives a secret link which usually expires within 1-12 months.

3.3 Making data interoperable

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

Partners will take into cosideration OpenAIRE guidelines for online data interoperability: https://guidelines.openaire.eu/en/latest/ and metadata standards according to the Metadata Standards Directory (Engineering, Physical Sciences & Mathematics and General Research Data) guidelines.

EVEREST project members will be using standard vocabulary for all data types. No data sets that require the development of novel specific standards and methodologies to facilitate interoperability have been generated in the project so far. All datasets generated are expected to map to commonly used ontologies.

3.4 Increase data re-use (throughout clarifying licences)

Template guiding questions:
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.)?
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?
Will the data produced in the project be useable by third parties, in particular after the end of the project?
Will the provenance of the data be thoroughly documented using the appropriate standards?
Describe all relevant data quality assurance processes.
Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

Licensing

Where possible EVEREST data will be licenced following the guidelines provided by CESSDA Training (Licensing your data) and Creative Commons Licence (https://creativecommons.org/choose/) or (http://ufal.github.io/public-license-selector/) by integrating the appropriate abbreviation into the shared file.
According to EVEREST Grant Agreement, data and software are owned by the beneficiary that generates them. Nevertheless, the owners of open results arising from the project are encouraged to release their work preferably under a Creative Commons Attribution 4.0 (CC-BY-4.0).

An embargo period (6-12 months depending on the journals) may be applied if the data are used in published articles in "Green" open access journals.

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

For datasets deposited on a public data repository (i.e. Zenodo) the access is unlimited.

Restrictions on re-use policy are applied for all protected data, whose re-use will be limited within the project partners.

Longevity

Data necessary to validate the results presented in the publication will be archived in a FAIR-compliant repository and be linked to the publication. Raw data not directly relevant to the publication will be stored securely and made available upon justified request. Data associated with public deliverables will be shared once the deliverable has been approved and accepted by the EC. For other public datasets not directly linked to a scientific publication or deliverable, such datasets will be made available upon assessment by relevant Consortium body that it is ready for publishing, and in the final month of the project at the latest.

All data necessary to validate the results presented in the publication will be archived in a FAIR-compliant repository. Raw data not directly relevant to the publication will be stored securely and made available upon justified request.

Open data can be reused in accordance with the Creative Commons licences. Data classified as confidential will as default is not reusable due to privacy concerns.

The public data will remain reusable via Zenodo for at least 20 years, according to Zenodo's general policies (http://about.zenodo.org/policies/).

Since no research data has been produced within the project to date, the specific question of its re-usability by third parties or usability period is not fully developed in the current version of the DMP document.

3.5 DMP quality and review

Quality evaluation and reporting processes are implemented throughout the entire project to assess both project outputs (data/products) and overall progress (refer to Deliverable 5.2 - Project Quality Plan). Key deliverables undergo an internal peer review to ensure high-quality standards. WP leaders must submit all documents to a designated internal reviewer from another partner for quality checks.

Project results and data will be continuously analyzed and collected throughout the project's duration. To support this, each partner will update the Communication Plan (Deliverable 4.3) once a year, documenting articles, papers, and scientific publications. A dedicated table (currently under development) will be used to record all research data and publications. WP leaders and publication authors are required to update this table periodically.

The Data Management Plan will be regularly updated to reflect any changes, including revisions to the online research data repository. This repository will include descriptions of datasets and research data as they are generated and collected.

3.6 Management of other research output

Beyond data management, beneficiaries should plan for the management of additional research outputs generated or reused during their projects. These outputs may be digital (e.g., software, workflows, protocols, models) or physical (e.g., new materials, antibodies,

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

reagents, samples). Beneficiaries should evaluate which FAIR data principles apply to these outputs and provide detailed plans for their management, sharing, and reuse.

4 Data security

__Template guiding questions__: 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)?
Will the data be safely stored in trusted repositories for long term preservation and curation?

4.1 Data access

To increase data security, the following measures will be implemented:

- Raw data will be stored in at least two separate locations to prevent loss. Digitalized data is only stored on the main institutional server with only selected data being deposited into EVEREST TEAMS.
- Encryption will be applied if deemed necessary by the researchers involved.
- The use of USB flash drives will be minimized.
- Files will be labeled systematically to maintain consistency within the final dataset.

All project-related deliverables and data will be stored and shared exclusively within a dedicated Microsoft TEAMS repository (certified repository with appropriate security protocols), accessible only to the project consortium. Initially, access to the repository, where datasets and metadata are organized, will be limited to Consortium Partners. Data protection will be ensured through established procedures and suitable technologies, such as the HTTPS protocol for encrypted internet transactions, alongside compliance with European and global security standards (ISO, ITU, W3C, IETF, and ETSI).

Scientific publications and articles, and other shareable research outputs will be made available through Zenodo, its safety is guaranteed according to the product description (see https://zenodo.org/features).

4.2 Data preservation

Data preservation refers to data upkeep and maintenance to ensure that the integrity of the data is upheld in the future. This includes properly maintaining the data repository and data backups to ensure the long-term value of the data.

Data backups are expected to occur at least once a week, though this is subject to change based on the amount of data that will ultimately need to be backed up. Estimating the exact size of the data is difficult to accomplish at this time since the data that will be initially included in the project has not been fully identified yet (currently at M3 of the project). An initial inventory of up to 13 TB volume of data is identified by the Consortium in the Table 2. More exact estimation on the data size will be made closer to project's mid-term.

TEAMS project repository has no limits on the size of data that can be stored.

The EVEREST data can also be downloaded and copied to personal computers by any partner having access to the repository.

5 Ethical aspects

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

Template guiding questions:
Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?
These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).
Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

5.1 General Data Protection Regulation

Partners shall cooperate in order to enable one another to fulfil legal obligations arising under applicable data protection laws (*the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data* and/or relevant national data protection law applicable to said Partner) within the scope of the performance and administration of the project and of the EVEREST Consortium Agreement; and comply with regulations on intellectual property rights (IPR).

Access to the project repository and database will remain entirely anonymous. Users wishing to utilize the TEAMS platform for the EVEREST project must register by creating an account. During registration, users are required to provide an email address and password and consent to the processing of their email address for account-related communication and management purposes. Users may withdraw this consent at any time, which will result in the deletion of their account.

5.2 Sensitive data

Sensitive data is data that is either private or confidential and includes personal user data. The proper management of sensitive data is imperative to maintain the individual privacy and remain in compliance with both EU and international regulations.

In order to ensure sensitive data is properly managed, data that is considered sensitive should first be identified. Thus, the main ethical and privacy issues with sensitive data arise from ensuring the data remains private and that proper consent is obtained before the data is shared or published in any way.

When possible, collected response data will be anonymized so that it cannot be directly attributed to the responder (for example, by delineating a numeric code to an individual). In addition, data will be reported in aggregated forms to further prevent any firm or individual from being identified through their response. If anonymization is not possible, then the explicit permission will be received prior to the publication of sensitive data. In any case, sensitive data will always remain confidential.

To ensure proper communication with external audience, any person can subscribe for EVEREST Newsletter. The collected data represents a list of email address and names associated with them. The data is stored on Brevo's servers, which are GDPR compliant. More details can be found here: https://help.brevo.com/hc/en-us/sections/18503544961042-GDPR. At current stage of the project there are no plans to generate or re-use any other personal user data. If sensitive data generation is anticipated at later stage, the lead partner will take consultancy measures from their local data protection officer regarding the following steps for the data use, storage and access.

6 Resource allocation

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

__Template guiding questions__: 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.) ?
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)
Who will be responsible for data management in your project?
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)?

Costs

Generally, estimation of costs of data management activities is not straightforward, as many activities are an integrated with standard research process and data analysis. Ideally, it is necessary to estimate the time-cost related to data collection, data entry and transcription, data validation and documentation and the cost of preparing data for archiving and re-use. Those resources that include time and effort costs (i.e. data records preparation, maintenance of technical infrastructure, individual preparation effort needed to use the infrastructure etc.) are so-called non-monetary costs. Since the researchers and the management team are the main producers of data in EVEREST, all these costs are related to them. The consortium expects that monetary costs for FAIR data will be mainly related to Open Access publication of papers (research articles and reviews), combined with patent protection of exploitable results, and maintenance of hosting infrastructure (servers) for data storage. According to the EVEREST GA, Open Access publication related costs sum up to 6000 EUR as a lowest margin (publication fees alone).

Monetary costs related to Open Access data in HE are eligible for reimbursement under the conditions defined in the GA, but also other articles relevant for the cost category chosen. Each partner is responsible for managing the costs related to making data accessible to others beyond the consortium.

Some examples of costs for making data FAIR include:

- "Gold" Open access publication fees. The cost sharing, in case of multiple authors, shall be decided among the authors on a case-by-case basis;
- Project Website operation (to be defined by the 1st reporting period);
- Data archiving at Zenodo and its national homologues: free of charge
- Copyright licensing with Creative Commons: free of charge.

No immediate costs are anticipated for open data that is stored for long-term preservation. Additional details will be reported, as needed, in future versions of the DMP.

Regarding the question of long term raw data preservation, no specific arrangements has been done in the consortium yet, although, with a great degree of confidence we can state that raw data will be kept in available repositories at each institution. On later stages, the consortium will decide whether some raw and processed data can also be long term stored in the common storage space that will be allocated for the project shortly.

Responsibility

Data management activities concern the whole project and needs to be coordinated and monitored both at project and WP level. Data management is also linked to publication of project results and thus dissemination activities.

The overall responsibility for data management lies with the project coordinator GRS.

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.

The members of the Executive Committee (i.e. WPLs) are also responsible of the Data Management of EVEREST datasets and research data in agreement with each organization's internal process, outlined by their local Data Protection Officer (DPO).

The Coordinator (project data manager) and WPLs (WP data managers) are responsible for:

- Co-developing and implementing the data management plan and policy with project management.
- Monitoring data collection, publication activities, and deadlines.
- Ensuring open results (data/software) are deposited in compliant repositories and linked to the project.
- Writing, updating and uploading the DMP.
- Advising on publication paths (green/gold open access) and offering customized help and further guidance for publishing scientific publications.
- Verifying journal compliance with HE open data policies before manuscript submission.
- Ensuring green access publications are deposited in repositories.
- Making publication metadata available in the R&I Participant Portal and EVEREST website.
- Ensuring research data linked to publications is in repositories.
- Tracking embargo periods and notifying partners.
- Linking OpenAIRE publications with the Project.
- Describing data or publications (by means of appropriate metadata) per EVEREST policies and using project-provided templates/tools.

7 Conclusion

This document outlines the core principles and guidelines for data management within the EVEREST project. As a dynamic document, it will be revised and updated throughout the project's duration. In the future updates to the DMP we plan to address progress regarding the online research data repository, including the collection, sharing, and detailed description of datasets and research data as they are progressively generated and gathered.

---- End of the document ----

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Atomic Energy Community ('EC-Euratom'). Neither the European Union nor the granting authority can be held responsible for them.