EASI SMR

# D9.3 Data Management Plan


Author: EDF

*Contributions by all EASI-SMR partners*





Co-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







Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

# List of Figures



# List of Tables



# Document information

| | |
|---|---|
| **Grant Agreement Number** | n°101164810 |
| **Project Title** | Ensuring Assessment of Safety Innovations for SMR |
| **Project Acronym** | EASI-SMR |
| **Project Coordinator** | Nicolas Sobecki |
| **Project Duration** | 1 September 2024 – 31 August 2028 (48 months) |
| **Related Work Package** | WP9 |
| **Lead Organisation** | EDF |
| **Contributing Partner(s)** | All Partners |
| **Submission Date** | 28/02/2025 |
| **Dissemination Level** | Public |

# Disclaimer



# History of changes

| Date | Submitted by | Reviewed by | Version (Notes) |
|---|---|---|---|
| 31/01/2025 | External PMO | Nicolas SOBECKI (COO) | V0.1 First internal draft sent for review.<br>Figure 2 added<br>Section 4 reworked |
| 13/02/2025 | External PMO | All partners | Caveat added on p.4;<br>Other national repositories added (3.1);<br>Data upload on Zenodo clarified (3.2.2);<br>Access to the restricted data (in 3.2.4) is complemented by the role of each partner;<br>Purpose of each data hosting site clarified (4);<br>template guidelines removed |
| 28/02/2025 | External PMO | Nicolas SOBECKI (COO) | V1Final Version for submission is ready |

# Summary

This deliverable presents the first version of the EASI-SMR 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 document follows the model, provided by the EC[1].
The purpose of the DMP is to contribute to good data handling through indicating what research and other satellite 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

Caveat: The provisions of the GA and CA take precedence. None of the principles and instructions of the DMP should be interpreted in such a way that they contradict the GA and CA.

# Keywords

EASI-SMR, SMR, Data Management Plan, open data, research data, data repository, data storage, Zenodo, FAIR data

# Abbreviations, acronyms and definitions

| Acronym | Description |
|---|---|
| AB | Advisory Board |
| AE | Affiliated Entity |
| AP | Associated Partner |
| BEN | Beneficiary |
| CA | Consortium Agreement |
| COO | Coordinator |
| DoA | Description of Action |
| EC | European Commission |
| ExCom | Executive Committee |
| EUG | End User Group |
| GA | Grant Agreement/ |
| GeA | General Assembly |
| HE | Horizon Europe funding programme |
| LW-SMR | Light Water SMR |

[1] https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdfhttps://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf





Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

| | |
|---|---|
| PMO | Project Management Office |
| PMP | Project Management Plan |
| PQP | Project Quality Plan |
| RP | Reporting Period |
| WP | Work Package |
| WPL | Work Package Leader |

| Term | Description |
|---|---|
| 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 Charges (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), however, HE rules require the publication be available in online repository at the same time as article's peer-reviewed publication.[2] |
| Personal Data | Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which 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 |
| DOI | Digital Object Identifier: a unique identifier in a format of a alphanumeric string assigned to digital documents, such as journal articles, research papers, datasets, and reports, to provide a permanent link to their location on |



[2] Fact Sheet Open Access to scientific publications and research data: https://www.horizon-europe.gouv.fr/sites/default/files/2021-07/h2020---ipr-helpdesk---open-access-3953.pdf

| | |
|---|---|
| | the internet. DOIs help ensure long-term access to digital content, even if the URL changes. |

# 1. Introduction

This deliverable describes the data management life cycle for the data to be collected, processed and/or generated by the EASI-SMR 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:

- 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).

According to the EU's guidelines on the DMP, this document may be updated - if appropriate - during the project lifetime. 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 change log table with modifications summary.

# 2. Data summary

The EASI-SMR project's main objectives are directly linked to facilitate the use of LW-SMRs in energy transition and industry decarbonation in Europe by resolving the safety issues associated with its major innovations. The main activities, for which the dominant part of data is collected, generated or re-used within each WP (see category "Research and Scientific Data" in Table 1) are the following:

1. Work towards LW-SMR acceptability and licensing
2. Experimental test program
3. Code validation and scaling
4. Reliability of passive systems
5. Study of human and organizational factors
6. SG Mock-up by additive manufacturing techniques.

Meeting these objectives requires cooperative work of over 300 researchers from 38 partnering institutions across Europe and beyond. Each partner will collect or generate specific data, related to their tasks as per the GA. This first version of the DMP is the result of the early stage data inventory by the EASI-SMR partners. During the course of the Action we are to correct the information given below so that it better reflects the real state of 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. Some preliminary data summary is provided in the Table 1 below.
It contains responses to the following questions, addressed by all partners and coordinated by each WPL:

- 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

More information and guidelines to better understand how to properly address the points above could be found via the Digital Curation Centre[3] (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] https://www.dcc.ac.uk/guidance





Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

## Table 1. EASI-SMR Data Inventory Table

| | Data Purpose | Data Origin | Data Types | Data Formats | Existing data to be used or re-used | Estimated Data Size, that the project indends to generate or re-use | Data Utility- to whom can your data be useful outside the project? |
|---|---|---|---|---|---|---|---|
| Research & Sientific data | In WP1 collected data will be used for requirement descriptions, benchmark designs, waste management system issues, building stakeholder interaction, EPZ calculation, co-location of nuclear sites, integration to the electrical grid, damage and recovery actions determination | Calculations, code simulations, public data, data from design companies, literature, interviews, observations in environments, radiological survey data | Core design data, text, recordings, questionnaire responses, interview transcriptions | *.xlsx, *.doc, *.pdf, computational tools output | Data and information from ended or on-going EU projects (ELSMOR, HARMONISE, FORSAFE, TANDEM), available data on operational events, documentation on nuclear facilities and radiological survey data | <10GB | Regulators, licencees, vendors, utilities, TSOs, WMOs, civil organisations, municipalities |
| | In WP2 collected data will be used for code benchmarks for WP3 . Data will be used too for WP4 statistical methods application. | primary data generated by 9 experimental tests campaigns : COSAC, FHEASIK LDR50, GRADAC, ELSMOR II, IVR LOOP, PANDA, PRECISE, ALCINA, ECRINS<br>•Preliminary definition of tests and design components<br>•Basic design data of the facility, with production of matrix data and figure of merit for the benchmark preparation<br>•Detailed design data: experimental results for code benchmark in specific experimental facility. | •Experimental test facility and test matrix description<br>•Experimental test results for each facilities, each sensors and each test series as a function of time (temperature, pressure, flowrate...). | *.xlsx, *.doc, *.pdf<br>CAO file (.step) | Data and information from existing ELSMOR, PANDA, IVR LOOP and PRECISE test facilities | <40GB | Regulator, vendors, universities, TSOs, utilities |
| | In WP3 the data will be used to develop the code benchmark exercises. Data will be also used to develop a proof of feasibility of a large scale IETF. | Data generated by the development of code input-decks, and code simulations.Data generated, also, from the design of the IETF and related code input-deck development and simulations | Computer code input-deck files, computer code simulation results; other data ( e.g. immages, etc) | According to computational tools requirements and depending on computational tools outputs: ASCII *.csv, *.dat., *.xlsx, *.doc,. *.pdf, *.gng, *.jpeg, *.zip, *.txt, *.dwg, etc | The WP3 exercise are benchmarks against experimental data developed in the project. For the development of the proof of feasibility of a large scale IETF may be used data also from the ELSMOR E-SMR or from public literature data | <10 GB | Regulator, vendors, universities, TSOs, utilities |
| | In WP4 collected data will be used for estimation of the passive system reliability (failure probability), development of the PSA model and benchmark of the results. | Primary data: results of the T-H calculations, secondary data (for PSA model development): public available (IAEA, NRC, etc) | T-H codes calculations outputs, component reliability data (e.g. failure rates) | *.xlsx, *.doc, *.pdf | Data and information from existing ELSMOR facility, other public available resources (e.g. NRC or IAEA database for components reliability parameters) | <1 GB | Regulator, vendors, universities, TSOs |
| | In WP5 collected data will be used to get as close as possible to the future human activities involved in the manufacture/operating/maintenance of SMRs and to highlight the potential effects of SMR specificities on these activities | Literature, Interviews, Questionnaires, Activity observations in real and simulated environment | Texts (scientific articles, procedures, documentation relating to SMR design hypotheses, interview transcriptions), questionnaire responses, audio-video recordings | *.xlsx, *.doc, *.pdf, .mp4, .mov, .avi, etc. | Literature, documentation relating to SMR design hypothesis, other public available resources | <400GB | The raw data collected during the project, concerning human activities, are considered sensitive and cannot be shared. Some data may be shared (e.g. interviews), provided it is completely anonymized. |
| | In WP6 collected data will be used for produce additive manufacturing mock-ups of steam generators | Experimental data obtained in operation conditions; comparison to conventional materials. Test conditions and test matrix defined based on the previous experience in the NUCOBAM project | DESIGN - Design AM, PROCESS - Process planning for AM, BUILD - Part build, POST - Post-processing, QUAL - Qualification and certification | *.xlsx, *.doc, *.pdf, etc. | open deliverables from NUCOBAM project | < 5GB | Anyone from the component supply chain, regulators. utilities |
| | In WP7 data of two core designs (LDR Lite and PRATIC) is collected for the realization of all the tasks within the WP7 | LDR Lite data is from VTT while PRATIC data is from CEA | Data about the geometry, dimensions, materials and operation conditions of the two reactor designs. In addition, neutron physical data of the first and equilibrium core of both designs, about fuel management strategy of both designs. | *.xlsx, *.doc, *.pdf, etc. | none | < 5 GB | Industry, regulator, universities, TSOs, research centres |
| | In WP8 part of the collected and generated data is relevant to the Dissemination & Communication category below. | | | | | | |
| | In WP9 : all generated data is relevant to the MGMT & Legal data category below | | | | | | |
| 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 guidelines 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, *.doc, *.pdf, etc.<br>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<br>General public<br>Projects/Initiatives : EU funded projects |
| Dissem. & Commun. Data | ENEN: Data will be collected from WP1-7 for further utilisation during the project E&T activities planned in WP8<br><br>LGI: Data will be generated to support communication and dissemination of the project results and objectives | ENEN: research and scientific outcomes of WP1-7<br>LGI: Photos, progress reports, project's events, stock images | ENEN: WP1-7 reports, presentations, photos, diagrams<br>LGI: Partners' logos and promotional images/templates, Web digital content, hard copies | *.xlsx, *.doc, *.pdf, *.ppt, etc. visual/image formats<br>*.json | ENEN: involved organisations may add some content to supplement E&T materials with the research information in their organisation<br>LGI: No data from previous work will be used or re-used for the communication or dissemination purposes ofthis project | ENEN: < 500 MB<br><br>LGI: Data size < 500 MB | EU institutions & Regulators<br>General public<br>Specialized public: Research peers & Students |
| 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.<br>*The website does not use cookies.<br>**The project does not organize recruitment campaigns and does not generate or store candidates data<br>During the interviews/activity observations, WP5 will gather information such as surname, first name, institution and career details. | Subscribers contacts<br>Collaborators contacts<br>Attendees contacts (in case of external events)<br>Interviews & Activity observations in real and simulated environment | Email address/Names<br>Partner Entity<br>Role in the project<br>Interview transcriptions, Audio-video recordings | *.xlsx, *.doc, *.pdf, *.mp4, *.mov, *.avi, *.json etc. | No data from previous work will be used or re-used | Data size < 400 GB | By default any personal data will not be used outside the project.<br>In case any part of personal data is to be used publicly, it will either be anonymized or a prior usage rights authorization will be obtained from any concerned private body.<br>Collected data is GDPR compliant (see Section 5.2) |







Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by
Schweizerische Eidgenossenschaft Confédération suisse Confederazione Svizzera Confederaziun svizra
Swiss Confederation
Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

# 3. Data Management Policy: FAIR data

EASI-SMR'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 EASI-SMR and that are to be made available by the project consortium as open source, open science and open data.

## 3.1. Making data findable

This section will provide insight into how EASI-SMR intends to make it easier to find data collected or produced by the consortium. The way to proceed in order to achieve this goal is to describe properly the content of the data sets using appropriate tools, regulations and standards as described hereafter.

**Digital Object Identifier (DOI)**
All open data, publications, and open-source software generated within the EASI-SMR 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 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 EASI-SMR:

- 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, …),




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

- 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 EASI-SMR project will store its open data in Zenodo, an open-access online repository designed for research data: https://zenodo.org/communities/easi-smr. 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 EASI-SMR 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:

EASI-SMR, LWR-SMR, PWR, passive safety system, streamlined licensing, code-validation, reliability assessment, nuclear physics, additive manufacturing, thermal hydraulic test facility, boron free core, nuclear safety, nuclear reactor.

**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
- TESEO (https://www.educacion.gob.es/teseo/listarBusqueda.do) for Spanish PhD theses.
- Thesis in nuclear engineering prepared at the Jozef Stefan Institute (Slovenia): https://repozitorij.uni-lj.si/

**Naming conventions**
Files and folders at data repositories will be versioned and structured by using a name convention consisting as follow: FileType_EASI-SMR_[name of the document]_Vxy. FileExtension.
FileType can be:
D for Deliverable
DS for DataSet
F for File (generic, ex. images, table, document).

**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

This section will elaborate on consortium efforts to maximize openness and accessibility of research data generated within the EASI-SMR project. By default, research data and deliverables are shared within and beyond the consortium, ensuring broad dissemination to the scientific community and relevant stakeholders. Open access is facilitated through various approved platforms, including open access scientific publications, open data repositories, and the project's dedicated website. The vast majority of project deliverables will be made publicly available, reflecting EASI-SMR's dedication to transparency and knowledge sharing. While efforts are made to ensure unrestricted access, there is space for exceptions where legal, ethical, or intellectual property concerns may apply.

### 3.2.1. By default openly available data

In order to maximise the impact of EASI-SMR 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 EASI-SMR, we have strived to make as many deliverables public as possible: 74 out of 81 deliverables will be open for public access.
The EASI-SMR 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 EASI-SMR public datasets and deliverables (data access policy unrestricted) will be accessible by:

- EASI-SMR project web site: https://www.easi-smr.eu/
- Partners database
- OpenAIRE
- Zenodo and its national homologues for ORDP data and datasets
- Open access journals

All data deposited on Zenodo are accessible without restriction to the public.
Exceptions are the following:

- Copyright and permissions for re-using third-party datasets
- 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 models.
- Data that will be exploited commercially will not be made open.

### 3.2.2. Data Accessibility

EASI-SMR 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 curated and archived by the entity, which generated it..
Zenodo-hosted EASI-SMR community is now accessible via the following link: https://zenodo.org/communities/easi-smr and is managed by the Coordinator with the PMO support. Each Partner is to provide the files (publications, datasets, workshop materials etc), to be uploaded to Zenodo by the Coordinator/PMO.

**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 EASI-SMR and the related bibliographic metadata must be made available as open access and published on the project website as well as linked to the OpenAIRE portal (https://www.OpenAIRE.eu/) and the Funding & Tenders Portal (https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/results-innovation-support). To automate the process of linking scientific publications and related research data to OpenAIRE, the publication should be deposited in an OpenAIRE-compliant repository (such as Zenodo or Re3data, for example), 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




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

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 publications of the EASI-SMR project. Nevertheless, the journal 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 (AAM or VoR) 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, without an embargo period (immediate open access-HE requirement). To find 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.

A simplified schematics to depict various pathways to render publications openly accessible and HE compliant is given in Figure 1 below:




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

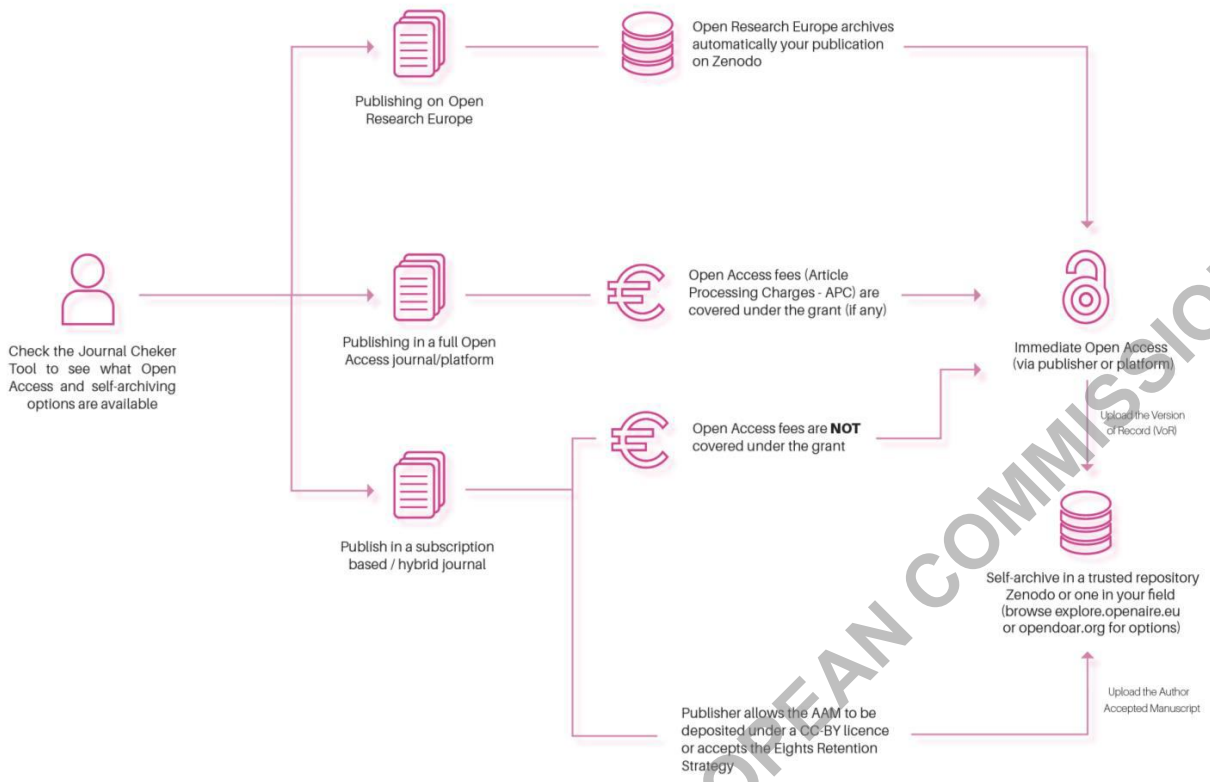


**Figure 1. Open Access schematics, image adapted from OpenAIRE webinar England & Malaguarnera2022.10.5281/zenodo.7324364**

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

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. This should be done in line with the chapter 9 "Access Rights" of the EASI-SMR Consortium Agreement. 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 of 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. The Partner or group of Partners owning the restricted data are responsible for installing and handling the controlled access to their data. 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




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

Zenodo repository", metadata for both open, closed, embargoed 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. Potential end-users 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

Partners will take into consideration 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.
EASI-SMR 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)

This section outlines the foreseen licensing practices that support compliance with FAIR principles, and archival policies that will guide the responsible sharing and preservation of project data, balancing open access with necessary restrictions on protected datasets.

**Licensing**
Where possible, EASI-SMR 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 EASI-SMR 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).
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 (in our case Zenodo) 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 the Executive Committee that it is ready for publishing, and in the final months 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 9.1 - Project Quality Plan and D9.2 – Project Management Plan). All deliverables undergo an internal peer review to ensure high-quality standards. Deliverable leaders must submit their deliverables to their WP leader and the coordinator for quality checks.
Project results and data will be continuously analysed and collected throughout the project duration. To support this, each partner will update the Communication & Dissemination Plan (Deliverable 8.5) 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 outputs

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, 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 and dedicated platforms

To increase data security, the following measures will be implemented:
generated data will be stored in a way resilient to the data carrier damage, loss or malfunctioning: on institutional servers of each partner and three online sites as described further in this section.




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

Encryption will be applied if deemed necessary by the researchers involved.
The use of USB flash drives will be minimized.
Files will be labelled systematically to maintain consistency within the final dataset.
As for the estimation of the exact final volume of generated data – it 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 in M6 of the project). An initial inventory of up to 1 Tb of data is identified by the Consortium in the Table 1. More exact estimation on the data size will be made closer to the project's mid-term.

Three data hosting sites will be used within the framework of this project: consortium internal platform Teams, project's website and the EU Zenodo platform for dissemination of the publications and results.
All key public data produced by the project will be stored on Zenodo and a selection of files will also be made available on the project website.
Internal project data will be stored on the Teams platform. It will not be accessible for individuals external to the project.
For each platform, we will outline the security and storage of the data hereafter:

**Teams**
Purpose: Teams collaboration space (ShareFolder) will host most of the working documents relevant to the project implementation, as well as project's outcomes such as deliverables, publications, technical slides and support documentation.
Accessibility: EASI-SMR Teams is accessible only to the project consortium.
Data security 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).
Frequency of backups: daily backups to a datacenter in France (*Linkwe* service provider).
The security standards you meet or apply: Security Information and Event Management (SIEM), as well as IDS and IPS systems and complies with GDPR rules for personal data management.
What happens to the data once the project is finished? The content of the Teams collaborative space remains available 1 year after the end of the project, then it is archived on a local server in Paris (except personal data which are destroyed).
Data longevity: after the project end the data is stored on a local server in Paris, replicated on another site geographically, without time limit.
Typical size of Teams repository at the start of the project is 30 Gb and can be extended to up to several hundreds of Gb according to the project needs.

**Zenodo**
Purpose: Zenodo will be mainly used to host scientific publications, datasets, and key public deliverables of the project.
Data Protection / Backup: Zenodo data is located in a CERN datacenter (European Union) and a daily incremental backup is installed (for more details on server management and backup see https://about.zenodo.org/infrastructure/).
Access security: access to Zenodo requires password authentication. Data security is guaranteed according to the platform description (see https://about.zenodo.org/infrastructure/).




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

Confidentiality of sensitive data: by its nature and design, Zenodo does not allow the storage of confidential data, the data is public and freely accessible. However, Zenodo allows users to upload files under closed access, which means that the 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.
Data longevity: Zenodo supports long-term preservation of your data deposits, as the repository is projected to be maintained for the lifetime of the host laboratory CERN, defined as at least the next twenty years (see section 3.4 above).

**EASI-SMR WEBSITE**
Purpose: the website will be used to host public deliverables.
Frequency of backups: daily backup, replicated over 7 days.
The location of your storage center: the storage center is in Germany at IONOS in an ISO certified European datacenter.
The security standards you meet or apply: the website and applications do not use cookies. The access to the website and application is via an SSL certificate (https).
The website remains active for 3 years after the end of the project without content updates but with continued maintenance in operational condition. Then it is destroyed and the domain name terminated.
Data longevity: the data is stored on a local server in Paris, replicated on another site geographically, without time limit.

# 5.Ethical aspects

## 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 EASI-SMR Consortium Agreement; and comply with regulations on intellectual property rights (IPR).
Users wishing to utilize the TEAMS platform for the EASI-SMR 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




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

arise from ensuring the data remains private and that proper consent is obtained before the data is shared or published in any way, as for example, in case of event pictures and promotional images depicting persons.
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 form 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. Lastly, personal or human-related data is not part of the core EASI-SMR's activities (as opposed to the fields of medicine, genetics, sociology etc.), which minimizes any possible risk of mishandling sensitive data.
To ensure proper communication with external audience, any person can subscribe for EASI-SMR 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

Effective data management is a critical aspect of the EASI-SMR project, encompassing both financial and non-monetary considerations. Estimating costs for data-related activities can be complex, as many tasks are integrated into standard research workflows. This section outlines the key cost elements associated with data collection, storage, and accessibility, as well as the responsibilities of different stakeholders in ensuring compliance with data management policies. It also details the financial aspects of making data FAIR and the roles of the consortium members in overseeing these processes.

**Costs**
Generally, the estimation of costs of data management activities is not straightforward, as many activities are 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 generators of data in EASI-SMR, 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 possibly maintenance of hosting infrastructure (servers) for data storage.




Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by

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. This point will be updated in the next version of the document, as the consortium will have more information on the number and cost of the "gold" open access publications ;
- 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 have been done in the consortium yet, although, a) with a great degree of confidence we can state that raw data will be kept in the internal repositories at each institution in line with the local data storage guidelines (more details to be provided in the next version of the document), and b) long-term storage requirements will also be met with the help of the one-stop Zenodo repository. At later stages, the consortium will decide whether and how some raw and processed data can also be long term stored in the common Teams space that is allocated for the project by the PMO.

**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 EDF.
The members of the Executive Committee (i.e. WPLs) are also responsible of the Data Management of EASI-SMR 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 EASI-SMR 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 EASI-SMR policies and using project-provided templates/tools.

# 7.Conclusion

This document outlines the core principles and guidelines for data management within the EASI-SMR project. As a dynamic document, it will be revised and updated throughout the project 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.





Co-funded by the European Union
UK Research and Innovation
The Research Council of Norway
Project funded by