



ESFR-SIMPLE

Research and Innovation Action (RIA)

This project has received funding from the Euratom
research and innovation programme 2021-2025 under
Grant Agreement No 101059543

Start date : 2022-10-01 Duration : 48 Months

Data Management Plan

Authors : Mr. Pierre SCIORA (CEA)

ESFR-SIMPLE - Contract Number: 101059543

Project officer: Cristina FERNANDEZ RAMOS

Document title	Data Management Plan
Author(s)	Mr. Pierre SCIORA
Number of pages	34
Document type	Deliverable
Work Package	WP10
Document number	D10.2
Issued by	CEA
Date of completion	2022-12-23 11:25:36
Dissemination level	Public

Summary

This is deliverable 'D10.2 Data Management Plan' of the ESFR-SIMPLE project. This document is structured in three distinct parts. The first part provides some basic contextualisation of a data management plan in Horizon Europe projects and provides key definitions necessary to understand the deliverable. The second part aims to explain the scope of ESFR-SIMPLE Data Management Plan, in other words, the purpose of data collection in the frame of ESFR-SIMPLE, the type, format and the origin of the data to be collected and the use and re-use of the data. The third part details the various actions ESFR-SIMPLE will implement to make its data findable, accessible, interoperable and available for re-use. According to the EU's guidelines regarding the DMP (European Commission, 2016), the document may be updated - if appropriate - during the project lifetime (in the form of deliverables). DMPs should, therefore, have a clear version number and include a timetable for updates.

Approval

Date	By
2022-12-23 11:26:14	Mr. Pierre SCIORA (CEA)
2022-12-23 11:30:32	Mr. Pierre SCIORA (CEA)

Table of contents

1. INTRODUCTION	4
1.1 Background of the Data Management Plan in Horizon Europe	4
1.2 Definition	4
1.3 Versions	5
2. Scope and lifecycle of the ESFR-SIMPLE data	6
2.1 General framework for data collection	6
2.1.1 WP1 – Advanced design and material studies for ESFR	6
2.1.2 WP2 – Use of AI in ESFR: innovative monitoring application	10
2.1.3 WP3 – Modelling for the simulation of the hybrid system behaviour	12
2.1.4 WP4 – – SMR safety studies	15
2.1.5 WP5 SMR flexibility enhancement	19
2.1.6 WP6 – Advanced components for passive safety	19
2.1.7 WP7 – Innovative monitoring	21
WP8 – Optimised fuel element	22
2.1.8 WP9 – Dissemination, Education & Training, social aspect & public perception	25
2.1.9 WP10 – Project coordination & management	26
2.2 Use and Re-Use of the data	26
3. FAIR Data Management in ESFR-SIMPLE	28
3.1 Making data findable	28
3.1.1 Storing the data with datasets	28
3.2 Making data openly accessible	29
3.2.1 Data licensing	29
3.2.2 Datasets that could be made openly accessible in ESFR-SIMPLE	30
3.2.3 Datasets to remain confidential	30
3.2.4 Data storage in ESFR-SIMPLE	31
3.3 Making data interoperable & increasing re-use	31
3.3.1 Making data interoperable	31
3.3.2 Restrictions for re-use	31
3.3.3 Archiving and preservation	32
4. Human resources	33



5. Ethical aspects 33
5.1 GDPR 33

List of figures

Figure 1: Knowledge Management – Definitions and hierarchy 5

List of tables

Table 1: Table specifying the content of a dataset 29
Table 2: Example of licenses 30

Summary

This is deliverable "D10.2 Data Management Plan" of the ESFR-SIMPLE project.

This document is structured in three distinct parts. The first part provides some basic contextualisation of a data management plan in Horizon Europe projects and provides key definitions necessary to understand the deliverable. The second part aims to explain the scope of ESFR-SIMPLE Data Management Plan, in other words, the purpose of data collection in the frame of ESFR-SIMPLE, the type, format and the origin of the data to be collected and the use and re-use of the data. The third part details the various actions ESFR-SIMPLE will implement to make its data findable, accessible, interoperable and available for re-use.

According to the EU's guidelines regarding the DMP (European Commission, 2016), the document may be updated - if appropriate - during the project lifetime (in the form of deliverables). DMPs should, therefore, have a clear version number and include a timetable for updates.

Keywords

Data, Knowledge, Information, Data Management



1. INTRODUCTION

1.1 Background of the Data Management Plan in Horizon Europe

This document is the ESFR-SIMPLE Data Management Plan (DMP), a deliverable that has been required by the European Commission. According to the European Commissionⁱ, Open access (OA) refers to the practice of providing online access to scientific information that is free of charge to the end-user and reusable. 'Scientific' refers to all academic disciplines. In the context of research and innovation, 'scientific information' can mean:

- peer-reviewed scientific research articles (published in scholarly journals);
- research data (data underlying publications, curated data and/or raw data).

The rationale is that data management is not a goal in itself but a key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse. It follows from this that projects must aim to improve and maximise access to and re-use of research data generated while balancing openness and protection of scientific information, commercialisation and Intellectual Property Rights, privacy concerns, etcⁱⁱ.

Data Management Plans (DMPs) are a key element of good data management. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on the data life cycle:

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

While open access to research data has become applicable by default in Horizon Europe, the Commission acknowledges that there could be good reasons to keep the research data confidential.

The ESFR-SIMPLE DMP was written in reference to WP10 on Project Coordination and Management.

1.2 Definition

Before explaining the consortium strategy in terms of Data Management, several terms must be defined:

- **Data**: Data refers to unstructured facts and figures, which are not organised in any way and which provide no further information regarding patterns, context, etc. For instance, data on production, demand, results from technical tests and so on, is unstructured data.
- **Information**: For data to become information, it must be contextualized, categorized, calculated and condensed. Information thus paints a bigger picture; it is data with relevance and purpose. It may convey a trend in the environment, or perhaps indicate a pattern of sales for a given period of time.
- **Knowledge**: Knowledge is closely linked to doing and implies know-how and understanding. The knowledge possessed by every individual is a product of his/her experience and encompasses the norms by which s/he evaluates new inputs from his/her surroundings. For instance, knowledge is related to the know-how acquired in R&D projects, commercial activities or the expertise that is inherent to each partnerⁱⁱⁱ.



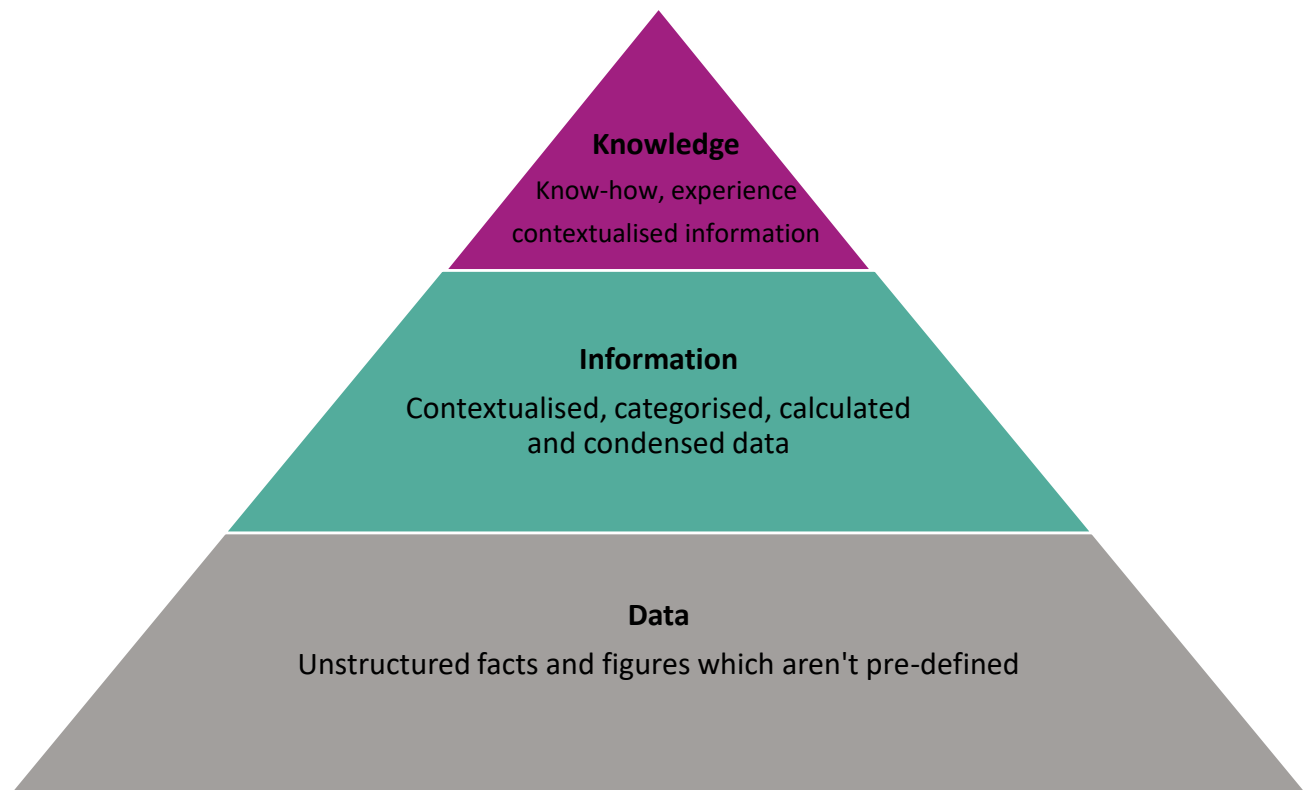


Figure 1: Knowledge Management – Definitions and hierarchy

This present DMP will mainly deal with how the data will be managed and will mention superficially the links with knowledge.

- Data codebook: A codebook is an essential document that informs the data user about the study, data file(s), variables, categories, etc., that make up a complete dataset. The codebook may include a dataset's record layout, list of variable names and labels, concepts, categories, cases, missing value codes, frequency counts, notes, universe statements, and so on^{iv}.
- Data set: a data set is a collection of data. Most commonly a data set corresponds to the contents of a single database table, or a single statistical data matrix, where every column of the table represents a particular variable. The data set lists values for each of the variables, such as height and weight of an object, for each member of the data set. The data set may comprise data for one or more members, corresponding to the number of rows^v.

1.3 Versions

According to the EU's guidelines regarding the DMP (European Commission, 2016), 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. If there are none, such an update must be made in time for the final review at the latest.

The DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses.

DMPs should, therefore, have a clear version number and include a timetable for updates.

2. Scope and lifecycle of the ESFR-SIMPLE data

2.1 General framework for data collection

In this section, the data to be collected in each WP of the ESFR-SIMPLE project will be presented and described in order to define the purpose of the collection as well as to previously define the type, format and origin.

2.1.1 WP1 – Advanced design and material studies for ESFR

The main objectives of WP1 are to optimize the secondary system design and to study pros and cons of using the new fuel type (metallic).

Name of the dataset	LV.IPUL.Closed.EMPDesign
Description of the dataset	<p>This dataset contains the design of electromagnetic pump design for the SMR with the main dimensions of the following components: rotor, pump channel, stator. It will be used by PSI and CEA to iterate on the design of the passive heat transfer system.</p> <p>Institutions involved in the design of SMR may be interested by this dataset. They can be used in publications on the ESFR-SMR design.</p>
Format/license	<p>The data will be available in the following format(s): *.ipt, *.iam, *.idw, *.dwg</p> <p>The license used for this dataset:</p> <p><input checked="" type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input type="checkbox"/>CC-BY-4.0 <input type="checkbox"/>ODbL</p> <p><input type="checkbox"/>Other, please specify:</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of IPUL

Name of the dataset	LV.IPUL.Closed.EMPTest
Description of the dataset	<p>This dataset contains experimental data of electromagnetic pump tests, with following parameters: temperature, flowrate, developed pressure difference, heat flux. It will be used by PSI and CEA to iterate on the design of the passive heat transfer system.</p> <p>Institutions involved in the design of SMR may be interested by this dataset. They can be used in publications on the ESFR-SMR design.</p>
Format/license	The data will be available in the following format(s): *.dat

	<p>The license used for this dataset:</p> <p><input checked="" type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p> <p><input type="checkbox"/>Other, please specify:</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of IPUL

Name of the dataset	DE.KIT.Closed.SIMMER.ESFR-SMR.ULOF-Results
Description of the dataset	This dataset contains SIMMER code inputs and selected simulation results of an ULOF in a ESFR-SMR model with the SIMMER code.
Format/license	<p>The data will be available in the following format(s): .txt and binary</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR

Name of the dataset	DE.KIT.Closed.SIMMER.ESFR-Metal-Results
Description of the dataset	This dataset contains SIMMER code inputs and selected simulation results of ULOF in ESFR with metal fuel.
Format/license	<p>The data will be available in the following format(s): .txt and binary</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR

Name of the dataset	DE.KIT.Closed.SIM-SFR.ESFR_MOX_fuel.P-Transient_Results
Description of the dataset	This dataset contains the selected simulation results of PLOF and PSBO transients for ESFR reactor, loaded with MOX fuel, obtained with SIM-SFR code. This is related to ESFR-SIMPLE Task 1.1 activities.
Format/license	<p>The data will be available in the following format(s): .xlsx</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR

Name of the dataset	DE.KIT.Closed.SIM-SFR.ESFR_Metallic_fuel.U-Transient_Results
Description of the dataset	This dataset contains the selected simulation results of ULOF and UTOP transients for ESFR reactor, loaded with Metallic fuel, obtained with SIM-SFR code. This is related to ESFR-SIMPLE Task 1.2 activities.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR

Name of the dataset	DE.KIT.Closed.SAS-SFR.ESFR_Metallic_fuel.U-Transient_Results
Description of the dataset	This dataset contains the selected simulation results of ULOF and UTOP transients for ESFR reactor, loaded with Metallic fuel, obtained with SAS-SFR code. This is related to ESFR-SIMPLE Task 1.2 activities.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR

Name of the dataset	CH.EPFL.Open.MetallicBenchmarckResultsOFFBEAT
Description of the dataset	This dataset contains the results obtained with the OFFBEAT code for the benchmark on metallic fuel behavior using SAS4A as a reference, in the framework of WP1 task 1.2
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input checked="" type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL <input type="checkbox"/> Other, please specify:
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories Zenodo The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	FR.EDF.Closed.EsfrCathareResults
Description of the dataset	This dataset contains the output of the simulation of several reactor transient on the commercial-size ESFR reactor with the CATHARE code. The estimation of the Steam Generator Heat transfer coefficient by dimensionless analysis will be included in the dataset, and should be used by partners for their own reactor simulation.
Format/license	The data will be available in the following format(s): .xlsx, .txt The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal EDF storage The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.HZDR.Closed.Serpent_ESFR-Metal_core_Fresh
Description of the dataset	ESFR core with metallic fuel: Serpent code inputs and selected simulation results for fresh core
Format/license	The data will be available in the following format(s): .txt, .xlsx, .m, and binary The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: ROssendorf DAta Repository (RODARE) The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.HZDR.Closed.Serpent_ESFR-Metal_core_Burnup
Description of the dataset	ESFR core with metallic fuel: Serpent code inputs and selected simulation results of fuel cycle analysis

Format/license	<p>The data will be available in the following format(s): .txt, .xlsx, .m, and binary</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: ROssendorf DAta Repository (RODARE)</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

2.1.2 WP2 – Use of AI in ESFR: innovative monitoring application

Considering the up-to-date state-of-the-art and the recent developments, the R&D activities suggested here for ESFR innovative monitoring have the following objectives to define specifications, use case and experimental conditions for the development of an AI (Artificial Intelligence) engine dedicated to preventive maintenance for a sodium fast reactor.

Name of the dataset	FR.CEA.Closed. Drawing of MECANA experimental loop
Description of the dataset	This dataset contains the design (CAD drawing) of the experimental part ; including different types of sensors and bubble generator
Format/license	<p>CAD drawing (Solidworks) - IDF</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	FR.CEA.Closed. Data issued from previous experiment and used for neural network training (phase 1)
Description of the dataset	This dataset data (measurement) contains data already acquired during past experiment
Format/license	Xls, national instruments, labview, txt The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	FR.CEA.Open. Specification of MECANA experimental loop
Description of the dataset	This dataset contains the specification of the experimental part ; including different types of sensors and bubble generator
Format/license	Microsoft World and PowerPoint The license used for this dataset: <input checked="" type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL <input type="checkbox"/> Other, please specify:
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Zenodo The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	FR.CEA.Open. Data issued from real time acquisition on MECANA experiment and used for neural network training (phase 2)
Description of the dataset	This dataset data (measurement) contains data acquired during the second phase of experiment

Format/license	<p>Xls, national instruments, labview, txt</p> <p>The license used for this dataset: <input checked="" type="checkbox"/> CC0 4.0 <input type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p> <p><input type="checkbox"/> Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Zenodo</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	DE.HZDR.Shared.AIBubbleClassificationImprovementAtHZDR
Description of the dataset	This dataset contains selected results of the evaluated high-resolution measurements and signals obtained by the various sensors, which are used as training data in task 2.3.
Format/license	<p>The data will be available in the following format(s): .xlsx, .csv</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Internal storage of HZDR</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

2.1.3 WP3 – Modelling for the simulation of the hybrid system behaviour

The main objectives of WP3 are summarized as follows:

- Formulate design criteria for ESFR-SMR in terms of safety, economics, and fuel cycle
- Propose a few ESFR-SMR designs that could potentially meet the specified criteria
- Assess fuel cycle and safety performance of the selected designs
- Compare ESFR and ESFR-SMR in terms of safety, economics, and fuel cycle.



Name of the dataset	FR.CEA.Open.SMRVesselDesign
Description of the dataset	<p>This dataset contains the design of potential vessels for the SMR with the main dimensions of the following components: vessel, intermediate heat exchanger, pumps and the core. It will be used by PSI and EDF to iterate on the design of the SMR core and vessel.</p> <p>Institutions involved in the design of SMR may be interested by this dataset. They can be used in publications on the ESFR-SMR design.</p> <p>These data are deduced from the dimensions of the ESFR-SMART reactor, supposing different core power and using the internal tool at CEA: COPERNIC.</p>
Format/license	<p>The data will be available in the following format(s): .xlsx</p> <p>The license used for this dataset: <input checked="" type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p> <p><input type="checkbox"/> Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Zenodo</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>
Name of the dataset	FR.EDF.Closed.SmrCoreDesign
Description of the dataset	<p>This dataset contains the design of the optimized SMR core(s), including power, fuel enrichment, number and size of the fuel assemblies, fuel pin size, and assembly axial structure.</p> <p>These data will be produced using the internal tool at EDF: SDDS</p> <p>It will be used by ESFR-SIMPLE partners for the reactor assessment.</p>
Format/license	<p>The data will be available in the following format(s): .xlsx</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p> <p><input type="checkbox"/> Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Internal EDF storage</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	CH.PSI.Shared.EsfrSmrCadDrawingsDataset
---------------------	---

Description of the dataset	This dataset contains the CAD drawings of small power ESFR version developed in WP3. There will be two subsets: the first subset is CAD drawings in dwg format and the second subset is the drawings in jpg format. The first subset will remain closed and available only for project partners, while the second subsets will be open inside and outside the project. In particular the second subset will be used in open publications and presentations.
Format/license	The data will be available in the following format(s): .dwg, jpg The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL <input type="checkbox"/> Other, please specify:
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: project Teams file storage. The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.HZDR.Closed.Serpent_ESFR-SMR_core_Fresh
Description of the dataset	ESFR-SMR: Serpent code inputs and selected simulation results for fresh core
Format/license	The data will be available in the following format(s): .txt, .xlsx, .m, and binary The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: ROssendorf DAta Repository (RODARE) The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.HZDR.Closed.Serpent_ESFR-SMR_core_Burnup
Description of the dataset	ESFR-SMR: Serpent code inputs and selected simulation results of fuel cycle analysis

Format/license	<p>The data will be available in the following format(s): .txt, .xlsx, .m, and binary</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: ROssendorf DAta Repository (RODARE)</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

2.1.4 WP4 – – SMR safety studies

The main objective of WP4 is the SMR safety analysis, based on the core and system design, as well as neutronics data coming from WP3.

Name of the dataset	FR.CEA.Shared.DesignMoltenFuelRelocation
Description of the dataset	<p>This dataset contains the simulation of a severe accident in the ESFR-SIMPLE reactor in order to evaluate the design safety measures on re-criticality prevention. Transient analyses will be performed by CEA with the SIMMER code, in order to investigate the performances of the available paths (transfer tubes) for molten fuel relocation.</p> <p>This dataset can be used in publications on the ESFR design. These data are deduced from the design of the ESFR-SMART core and reactor, that will be furnished beforehand.</p>
Format/license	<p>The data will be available in the following format(s): .txt files, binary files, .csv files, .xlsx files, .doc files and images (jpeg or png).</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p> <p><input type="checkbox"/>Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Zenodo and internal storage of CEA.</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	FR.CEA.Shared.MechanicalEnergyRelease
Description of the dataset	<p>This dataset contains the simulation of a severe accident in the ESFR-SIMPLE reactor in order to evaluate mechanical energy release potential. The effect of reactor size will be included by comparison with results which were obtained for ESFR-SMART.</p> <p>The investigation of the thermal-to-mechanical energy conversion following a hypothetical power excursion will be performed by CEA with the SIMMER code.</p> <p>This dataset can be used in publications on the ESFR design. These data are deduced from the transient configurations selected in Task 4.3.</p>
Format/license	<p>The data will be available in the following format(s): .txt files, binary files, .csv files, .xlsx files, .doc files and images (jpeg or png).</p> <p>The license used for this dataset: <input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL <input type="checkbox"/>Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Zenodo and internal storage of CEA.</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	DE.KIT.Closed.SIM-SFR.ESFR-SMR.P-Transient_Results
Description of the dataset	<p>This dataset contains the selected simulation results of PLOF and PSBO transients for ESFR-SMR reactor, obtained with SIM-SFR code. This is related to ESFR-SIMPLE Task 4.1 activities.</p>
Format/license	<p>The data will be available in the following format(s): .xlsx</p> <p>The license used for this dataset: <input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	DE.KIT.Closed.SIM-SFR.ESFR-SMR.U-Transient_Results
Description of the dataset	This dataset contains the selected simulation results of ULOF and UTOP transients for ESFR-SMR reactor, obtained with SIM-SFR code. This is related to ESFR-SIMPLE Task 4.2 activities.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.KIT.Closed.SAS-SFR.ESFR-SMR.P-Transient_Results
Description of the dataset	This dataset contains the selected simulation results of PLOF and PSBO transients for ESFR-SMR reactor, obtained with SAS-SFR code. This is related to ESFR-SIMPLE Task 4.1 activities.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.KIT.Closed.SAS-SFR.ESFR-SMR.U-Transient_Results
Description of the dataset	This dataset contains the selected simulation results of ULOF and UTOP transients for ESFR-SMR reactor, obtained with SAS-SFR code. This is related to ESFR-SIMPLE Task 4.2 activities.

Format/license	<p>The data will be available in the following format(s): .xlsx</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	FR.EDF.Closed.SmrSimmerUlofResults
Description of the dataset	This dataset contains the output of the simulation of an ULOF on the SMR with the SIMMER code.
Format/license	<p>The data will be available in the following format(s): .txt</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Internal EDF storage</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	DE.HZDR.Closed.ATHLET_ESFR-SMR_ULOF
Description of the dataset	ATHLET code inputs and selected simulation results of ULOF in ESFR-SMR with the ATHLET code.
Format/license	<p>The data will be available in the following format(s): .txt, .xlsx, and binary</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: ROssendorf DAta Repository (RODARE)</p> <p>The duration of the preservation will be: N/A</p>

	Foreseen costs of the preservation: N/A
	Means to cover preservation costs: N/A
	If the dataset contains personal data, will it be anonymised? N/A

2.1.5 WP5 SMR flexibility enhancement

The objective of this WP is to identify, and to preliminary design and demonstrate the additional benefits of a flexible SMR-SFR based on a TES coupled to the secondary loop..

Name of the dataset	SK.STUBA.Closed.SimuTesFlexibility
Description of the dataset	<p>This dataset contains the design of mechanical and electrical parameters for turbine, generator, their control systems, and electrical protections setting. Dataset can be used by TSOs (transmission system operators) to calculate Defense plans in their control area considering the assessment of the operational safety in the power system. Institutions involved in the power system operation may be interested by this dataset. Dataset can be used in publications on the TES flexibility quantification regarding to increasing the share of renewable energy sources in the power system.</p> <p>These data are deduced from the outputs of previous Tasks in WP5 focusing on the TES design coupled to the secondary loop.</p>
Format/license	<p>The data will be available in the following format(s): .doc</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL</p> <p><input type="checkbox"/> Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Zenodo</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

2.1.6 WP6 – Advanced components for passive safety

The main objectives of WP6 are develop and assess critical components related with reactor passive safety.



Name of the dataset	DE.KIT.Closed.LIVE-CC.Experiments
Description of the dataset	This dataset contains the selected experimental results of LIVE-CC. This is related to ESFR-SIMPLE Subtask 6.2.3 activities.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-ITES The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	FR.CEA.Shared.Design & experimental data
Description of the dataset	This dataset contains : <ul style="list-style-type: none"> • From design studies : CAO and report for thermomechanical studies • From experimental program : data (temperatures, expansion, ...) Tools used for these studies will be : ANSYS, COMSOL, CAST3M, RCC MRx and CEA tools. This dataset can be used in publications on the ESFR design.
Format/license	The data will be available in the following format(s): .txt files, binary files, .csv files, .xlsx files, .doc files, .step and images (jpeg or png). The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

2.1.7 WP7 – Innovative monitoring

Main objective is to develop and test various measuring techniques for the detection and characterization of gas bubbles in sodium.

Name of the dataset	DE.KIT.Closed.KARIFA.Sodium-boiling-experiments
Description of the dataset	This dataset contains the selected experimental results of KARIFA. This is related to ESFR-SIMPLE Task 7.5 activities.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of KIT-INR The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	DE.HZDR.Shared.Bubble measurements at HZDR
Description of the dataset	This dataset contains selected results on the motion of gas bubbles in the low-temperature liquid metal as well as in liquid sodium as planned in tasks 7.1 and 7.2.
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of HZDR The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

Name of the dataset	FR.CEA.Shared.Bubble measurements at CEA and bubble detection within secondary loop
Description of the dataset	This dataset contains selected results on bubble measurements at the CEA MECANA facility as well as the results on the secondary loop bubble detection and innovative steam generator design as planned in Tasks 7.3 and 7.4, respectively.
Format/license	The data will be available in the following format(s): CSV (concerning bubble detection) and STEP (for drawings) The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storages of CEA and HZDR The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

WP8 – Optimised fuel element

Objectives

- a parametric study of the behaviour of MOX pins during the transients considered in SFR (UTOP and ULOF). The objective is to assess the impact of each of the parameters (pin design, irradiation conditions, characteristics of the transients). This will orient the design choices in support of the ESFR-SMR.
- the fuel characteristics will be evaluated through their impact on the fuel properties at very high temperature (expected during transients): electronic and radiative contribution to thermal conductivity, effect of porosity on thermal properties to better control the fuel uncertainties.

Name of the dataset	IT_POLIMI_Open_Pomice
Description of the dataset	This dataset contains software capable of geometrically describing random heterogeneous materials. It does not contain any information related to specific materials, either in general or in specific instances.
Format/license	The data will be available in the following format(s): the format is a collection of text files. The license used for this dataset: <input type="checkbox"/> CC0 <input checked="" type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL

	<input type="checkbox"/> Other, please specify: MIT license
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: The data will be available on Github and Zenodo, with PID.</p> <p>The duration of the preservation will be: Indefinite. Foreseen costs of the preservation: None. If the dataset contains personal data, will it be anonymised? No personal data included.</p>

Name of the dataset	IT.ENEAClosed.SimulationBenchmark
Description of the dataset	This dataset contains the simulation inputs and results for the Benchmark exercise. Including the parametric study with calculations on 3 pin designs under 2 transients (ULOF, UTOP) with an evaluation of the margin to melt and risk of clad failure. The calculations are performed with the TRANSURANUS fuel performance code.
Format/license	<p>The data will be available in the following format(s): .txt The license used for this dataset:</p> <p><input checked="" type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p> <p><input type="checkbox"/> Other, please specify:</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of ENEA

Name of the dataset	FR.CEA.Closed.GerminalULOFResults
Description of the dataset	This dataset contains the results of the simulation of an ULOF with the GERMINAL code for the task 8.1.
Format/license	<p>The data will be available in the following format(s): .txt</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p>
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA

Name of the dataset	FR.CEA.Closed.GerminalUTOPResults
Description of the dataset	This dataset contains the results of the simulation of an UTOP with the GERMINAL code for the task 8.1.
Format /license	<p>The data will be available in the following format(s): .txt</p> <p>The license used for this dataset:</p> <p><input type="checkbox"/>CC0 <input type="checkbox"/>PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/>ODbL</p>

Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA
------------------------	--

Name of the dataset	FR.CEA.Closed.ExperimentalMeasurements
Description of the dataset	This dataset contains all characterisations and properties measurements of MOX fuels for the task 8.2.2.
Format /license	The data will be available in the following format(s): .txt, .png The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA

Name of the dataset	FR.CEA.Closed.AtomisticSimulation
Description of the dataset	This dataset contains the output of the simulation with atomistic tools for the task 8.2.3.
Format /license	The data will be available in the following format(s): .txt The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of CEA

Name of the dataset	CH.EPFL.Open.BenchmarkResultsOFFBEAT
Description of the dataset	This dataset contains the results obtained with the OFFBEAT code for the benchmark transient cases (UTOP & ULOF) identified in WP8
Format/license	The data will be available in the following format(s): .xlsx The license used for this dataset: <input checked="" type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL <input type="checkbox"/> Other, please specify:
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Zenodo The duration of the preservation will be: N/A Foreseen costs of the preservation: N/A Means to cover preservation costs: N/A If the dataset contains personal data, will it be anonymised? N/A

2.1.8 WP9 – Dissemination, Education & Training, social aspect & public perception

The main objectives are: 1. Develop new SFR computer simulator specifically for education and training purposes. 2. Compile, review and adapt a series of computational benchmarks to be used for education and training of young professionals on best practices and techniques of SFR modelling and design. 3. Organise and conduct a series of SFR technology-focused workshops and a Summer School to ensure information exchange and disseminate knowledge generated in the project. 4. Share the findings of ESFR-SIMPLE project with international organisations such as GIF. 5. Engage with general public through social media and other channels to disseminate knowledge and ideas created in the project. 6. Understand the public and stakeholder perceptions and needs for advanced reactor technology to inform the reactor design community decisions.

Name of the dataset	SK.STUBA.Open.Benchmarks_E&T
Description of the dataset	<p>This dataset contains a compilation of computational benchmarks selected by the project experts and adapted specifically for training purposes, focusing on skills and best practices of modeling SFRs. These include neutronics core modeling and transient, multi-physics simulation of the entire primary system loop.</p> <p>Institutions involved in the R&D of SMR, and E&T in nuclear engineering may be interested in this dataset. They can be used in publications on the ESFR-SMR design.</p>
Format/license	<p>The data will be available in the following format(s): .xlsx</p> <p>The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input type="checkbox"/> CC-BY-4.0 <input checked="" type="checkbox"/> ODbL</p> <p><input type="checkbox"/> Other, please specify:</p>
Archiving/preservation	<p>The data will be made available through the following platform(s) and/or repositories: Zenodo</p> <p>The duration of the preservation will be: N/A</p> <p>Foreseen costs of the preservation: N/A</p> <p>Means to cover preservation costs: N/A</p> <p>If the dataset contains personal data, will it be anonymised? N/A</p>

Name of the dataset	BE.LiegeUniv.Closed.projectpartners
Description of the dataset	<p>This dataset will contain the qualitative data (mostly text data & transcriptions) collected through several semi-structured interviews, observations, and a qualitative Delphi survey conducted with the ESFR-Simple project partners.</p>

	The data will be qualitatively analysed as part of the study on the social aspects of ESFR. This data will be used in a deliverable (D9.7) and in publications on the social aspects of ESFR.
Format/license	The data will be available in the following format(s): .csv files, .xlsx files, .docx files, and images (jpeg or png). The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of the Liège university. The qualitative data will be pseudonymized.

Name of the dataset	BE.LiegeUniv.closed.concernedpublics
Description of the dataset	This dataset will contain the qualitative data (mostly text data & transcriptions) collected through several semi-structured interviews, observations, and focus-groups conducted with members of concerned communities in France and in the UK. The data will be qualitatively analysed as part of the study on the social aspects of ESFR. This data will be used in a deliverable (D9.7) and in publications on the social aspects of ESFR.
Format/license	The data will be available in the following format(s): .csv files, .xlsx files, .docx files, and images (jpeg or png). The license used for this dataset: <input type="checkbox"/> CC0 <input type="checkbox"/> PDDL <input checked="" type="checkbox"/> CC-BY-4.0 <input type="checkbox"/> ODbL
Archiving/preservation	The data will be made available through the following platform(s) and/or repositories: Internal storage of the Liège university. The qualitative data will be pseudonymized.

2.1.9 WP10 – Project coordination & management

The main objective of WP10 is to carry out an effective technical, scientific, legal, financial and administrative coordination, establishing the mechanisms and management procedures to that end.

No data collection is foreseen within this WP.

2.2 Use and Re-Use of the data

The data collected and generated by the consortium will be useful to the development of further activities related to other WPs within the project, to specific end-users, plant designers, engineering companies and EU regulators.



As part of WP9, data will be re-usable only within the consortium, and for ESFR-SIMPLE related activities, such as newsletter dissemination, posters presentation, invitation to events or specific news on the website. WP9 will also handle exchange of data related to stakeholder exchanges.

The qualitative analysis of the public acceptance of the plant design would be reused in the framework of designing and implementing new nuclear plants.



3. FAIR Data Management in ESFR-SIMPLE

In compliance with applicable rules, every Horizon Europe project is required to draft a DMP in order to make the data Findable (1), Accessible (2), Interoperable (3) and available for Re-use (4) (FAIR principles).

3.1 Making data findable

This section will provide insight into how ESFR-SIMPLE 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.

3.1.1 Storing the data with datasets

- **Name and Reference code of datasets**

In order to imbue the names of datasets with easily identifiable meaning that conveys important information, the following naming convention shall apply:

CountryCode.DataOwner.Openness.Title

CountryCode: this string identifies the country to which the data pertains/where the data was collected using the ISO 3166 Alpha-2 coding system.

DataOwner: this string identifies the project partner in ESFR-SIMPLE that is associated with the dataset (data collector/custodian) using the official abbreviated partner names.

Openness: this string determines whether a given dataset is intended to be shared with the public as Open Data. It may take the following values:

1. **Open: can be accessed, used and shared by anyone without limitations, accessible on the internet in a machine-readable format, free of restrictions on use in its licensing)**
2. **Shared: available to use, but not under an open data license. Restrictions on its use or reproduction may apply (limited to a given group of people or organisations, may not be reproduced without authorisation, etc.)**
3. **Closed: can only be accessed by its subject, owner or holder**

Title: a short and descriptive string to identify the contents of the data

Using these strings, the name of a dataset would look like this:

FR.LGI.Open.CommuteHouseholdSurvey

A dataset with this name would describe a household survey on commuting preferences conducted in France and curated by LGI.

- **Description of the data**

Each data set that will be collected, processed or generated within the project will be accompanied by a brief table description. The following detailed information sheet will be produced for every dataset to be produced/collected/curated in the project.



Name of the data set^{vi}	Complete title of the data set
Description	<ul style="list-style-type: none"> • A brief, easy to understand description of what the dataset contains and what it will be used for in the project • A list of institutions to whom the data set could be useful outside the project • Whether the dataset has been/will be used for a scientific publication (if yes, brief details about the content and journal) <p>If the dataset is collected, a brief description of its origin and how it was collected will be provided</p>
Media Type	The physical medium of the content representation, e.g., video, image, text, numerical data, n-grams, etc.
Language(s)	The language(s) of the resource content
Use & re-use	Foreseen use of the resource for which it has been produced
Size	Size of the resource with regard to a specific size unit measurement in the form of a number
Format/license	The format in which the data will be available (e.g. .xls, .csv, .txt) will be provided. The license to be used will also be provided.
Version Number	Specify the version number of the document

Table 1: Table specifying the content of a dataset

If a dataset is directly collected, the origin of the data set will also be provided.

3.2 Making data openly accessible

Open accessibility of the data is the second key aspect for making data FAIR. This section will describe the type of data to be made available, its location and the procedure to obtain it.

Several degrees of accessibility are identified below, including both open access and restricted access.

3.2.1 Data licensing

Data licensing standards are used to layout the openness of data sets in concrete terms. There are many types of licenses to choose from, and this document will not cover them in depth. The table below provides a summary of common data licenses that will be considered for use in the project (based on definitions from opendefinition.org):

Name	Domain	Attribution	Share-alike*	Notes
Creative Commons CCZero (CC0)	Content, data	N	N	All rights (including those of attribution) waived
Open Data Commons Public Domain Dedication and Licence (PDDL)	Data	N	N	All rights (including those of attribution) waived

Creative Commons Attribution 4.0 (CC-BY-4.0)	Content, data	Y	N	Credit must be given, a link to the license must be provided, changes made must be indicated. If these terms are not followed, license may be revoked
Open Data Commons Open Database License (ODbL)	Data	Y	Y	Credit must be given, share-alike must be assured, data may be redistributed using DRM as long as a DRM-free version is also released

Table 2: Example of licenses

*Share-alike is the requirement that any materials created using the given dataset must be redistributed under the same license

3.2.2 Datasets that could be made openly accessible in ESFR-SIMPLE

Some of the **datasets in ESFR SIMPLE will be set as open. Here are the list of the open datasets:**

- CH.EPFL.Open.MetallicBenchmarkResultsOFFBEAT
- FR.CEA.Open. Specification of MECANA experimental loop
- FR.CEA.Open. Data issued from real time acquisition on MECANA experiment and used for neural network training (phase 2)
- DE.HZDR.Shared.AIBubbleClassificationImprovementAtHZDR
- FR.CEA.Open.SMRVesselDesign
- CH.PSI.Shared.EsfrSmrCadDrawingsDataset
- FR.CEA.Shared.DesignMoltenFuelRelocation
- FR.CEA.Shared.MechanicalEnergyRelease
- FR.CEA.Shared.Design & experimental data
- DE.HZDR.Shared.Bubble measurements at HZDR
- FR.CEA.Shared.Bubble measurements at CEA and bubble detection within secondary loop
- IT_POLIMI_Open_Pomice
- CH.EPFL.Open.BenchmarkResultsOFFBEAT
- SK.STUBA.Open.Benchmarks_E&T

3.2.3 Datasets to remain confidential

Due to the criticality of the project and given the potential misuse which could be done with the data, datasets of ESFR will have to remain confidential:

- LV.IPUL.Closed.EMPDesign
- LV.IPUL.Closed.EMPTest
- DE.KIT.Closed.SIMMER.ESFR-SMR.ULOF-Results
- KIT.Closed.SIMMER.ESFR-Metal-Results
- DE.KIT.Closed.SIM-SFR.ESFR_MOX_fuel.P-Transient_Results
- DE.KIT.Closed.SIM-SFR.ESFR_Metallic_fuel.U-Transient_Results
- DE.KIT.Closed.SAS-SFR.ESFR_Metallic_fuel.U-Transient_Results
- FR.EDF.Closed.EsfrCathareResults
- DE.HZDR.Closed.Serpent_ESFR-Metal_core_Fresh

- DE.HZDR.Closed.Serpent_ESFR-Metal_core_Burnup
- FR.CEA.Closed. Drawing of MECANA experimental loop
- FR.CEA.Closed. Data issued from previous experiment and used for neural network training (phase 1)
- FR.EDF.Closed.SmrCoreDesign
- DE.HZDR.Closed.Serpent_ESFR-SMR_core_Fresh
- DE.HZDR.Closed.Serpent_ESFR-SMR_core_Burnup
- DE.KIT.Closed.SIM-SFR.ESFR-SMR.P-Transient_Results
- DE.KIT.Closed.SIM-SFR.ESFR-SMR.U-Transient_Results
- DE.KIT.Closed.SAS-SFR.ESFR-SMR.P-Transient_Results
- DE.KIT.Closed.SAS-SFR.ESFR-SMR.U-Transient_Results
- FR.EDF.Closed.SmrSimmerUlofResults
- DE.HZDR.Closed.ATHLET_ESFR-SMR_ULOF
- SK.STUBA.Closed.SimuTesFlexibility
- DE.KIT.Closed.LIVE-CC.Experiments
- DE.KIT.Closed.KARIFA.Sodium-boiling-experiments
- IT.ENEA.Closed.SimulationBenchmark
- FR.CEA.Closed.GerminalULOFResults
- FR.CEA.Closed.GerminalUTOPResults
- FR.CEA.Closed.ExperimentalMeasurements
- FR.CEA.Closed.AtomisticSimulation
- BE.LiegeUniv.Closed.projectpartners
- BE.LiegeUniv.closed.concernedpublics

3.2.4 Data storage in ESFR-SIMPLE

After collection, data will be generally organised in Excel files and Word documents.

ESFR-SIMPLE will use Zenodo to systematically publish open data, open access presentations and public deliverables in order to maximise re-use and promote the project results. Prior to any upload, open publications on Zenodo will have to be approved by the Steering committee of ESFR-SIMPLE. If requested, LGI may provide its support and advice to the partners prior to the publication on Zenodo.

Storing data on Zenodo is free of charge and has no expiry date.

3.3 Making data interoperable & increasing re-use

3.3.1 Making data interoperable

As described on 3.1, Standard vocabulary may be used on a case-by-case basis to make the data interoperable between researchers, institutions, organisations, countries, etc.

- A list of acronyms and/or abbreviations will be provided at the beginning of every report
- Data will be stored using file formats in widespread use to maximise interoperability between software solutions, operating systems, etc.
- For surveys, standard definitions for entities such as trips, trip chains etc. will be adopted.

3.3.2 Restrictions for re-use

ESFR-SIMPLE will be compliant with the General Data Protection Regulation. To allow re-use, respect privacy and avoid loss of research data, two different techniques could be used to disseminate its data while abiding by regulations on privacy.



1) Anonymization of dataⁱⁱ

"Anonymization" of data means processing it with the aim of irreversibly preventing the identification of the individual to whom it relates. Data can be considered anonymised when it does not allow identification of the individuals it is related to, and no individuals can be identified from the data by any further processing of that data or by processing it together with other information which is available or likely to be available.

There are different anonymization techniques. Here are the two most relevant:

- Generalisation: generalising data means removing its specificity. For example, in the case of a table containing household income levels, with 4 figures mentioned: \$164,000, \$58,543, \$90,893, and \$232,234. One way of generalising this numbers would be to write that the values are "more than \$150,000, less than \$60,000, between \$90,000 and \$100,000, and more than \$225,000" respectively. Essentially it means taking exact figures, establishing a baseline category, and then obfuscating the data by assigning it to one of the categories in order to remove any sense of specificity from it.
- K-anonymity; A release of data is said to have the *k*-anonymity property if the information for each person contained in the release cannot be distinguished from the other individuals whose information also appear in the release. For instance, in a table composed of six attributes (Name, Age, Gender, State of Domicile, Religion and Disease), removing the name and the religion column while generalising the age is a way to effectively k-anonymise the data.

Other techniques, such as "masking" or "pseudonymisation", which are aimed solely at removing certain identifiers, may also play a role in reducing the risk of identification. In many cases, these techniques work best when used together.

2) Pseudonymisation of data

"Pseudonymisation" of data means replacing any identifying characteristics of data with a pseudonym, or, in other words, a value which does not allow the data subject to be directly identified.

Although pseudonymisation has many uses, it should be distinguished from anonymization, as it only provides limited protection for the identity of data subjects in many cases as it still allows identification using indirect means. Where a pseudonym is used, it is possible to identify the data subject by analysing the underlying or related data.

3.3.3 Archiving and preservation

It is of utmost importance for ESFR-SIMPLE to keep the data available for partners after the end of the project.

As already mentioned in 3.2, to ensure medium-term preservation of the datasets, anonymised data will be publicly shared by the consortium will be stored on Zenodo, which is a multi-functional open platform recognised by OpenAIRE and the European Commission.

4. Human resources

Every Work Package Leader in ESFR-SIMPLE will be responsible for the data management within its own Work-Package. LGI is responsible for authoring the present Data Management Plan in collaboration with project partners, and to make updates to it during the project as necessary.

5. Ethical aspects

5.1 GDPR

This Data Management Plan (DMP) was drafted and updated taking into account the General Data Protection Rules (GDPR) for the collection, storage and re-use of the data, in line with the following general principles.

Personal data shall be:

1. processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');
2. collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');
3. adequate, relevant and limited to what is necessary for relation to the purposes for which they are processed ('data minimisation');
4. accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy');
5. kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation');
6. processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality')^{viii}.

Bibliography

- ⁱ European Commission. Directorate-General for Research & Innovation (March 2017). HORIZON EUROPE Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Retrieved from:
http://ec.europa.eu/research/participants/data/ref/Horizon Europe/grants_manual/hi/oa_pilot/Horizon Europe-hi-oa-pilot-guide_en.pdf
- ⁱⁱ European Commission. Directorate-General for Research & Innovation. (July 2016). HORIZON EUROPE Programme – Guidelines on FAIR Data Management in Horizon 2020. Retrieved from:
http://ec.europa.eu/research/participants/data/ref/Horizon Europe/grants_manual/hi/oa_pilot/Horizon Europe-hi-oa-data-mgt_en.pdf
- ⁱⁱⁱ Alan Frost. Defining Knowledge, Information, Data. (2017). Retrieved from: <http://www.knowledge-management-tools.net/knowledge-information-data.html>
- ^{iv} DDI Alliance. Create a codebook. Retrieved from: <http://www.ddialliance.org/training/getting-started-new-content/create-a-codebook>
- ^v Rajit Dasgupta. 19 Free Public Data Sets for your first data science project. Springboard.com. (October 2015). Retrieved from : <https://www.springboard.com/blog/free-public-data-sets-data-science-project/>
- ^{vi} Maria Koutsombogera & Stelios Piperidis (Athena RC). D3.1 Data Management Plan. Cracker, Cracking the language barrier. (June 2015). Retrieved from:
<https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5a0203500&appId=PPGMS>
- ^{vii} Data Protection Commissioner. Anonymization and Pseudonymization. Retrieved from:
<https://www.dataprotection.ie/docs/Anonymisation-and-pseudonymisation/1594.htm>
- ^{viii} Intersoft Consulting. GDPR. Article 5. (2018) Retrieved from : <https://gdpr-info.eu/art-5-gdpr/>

