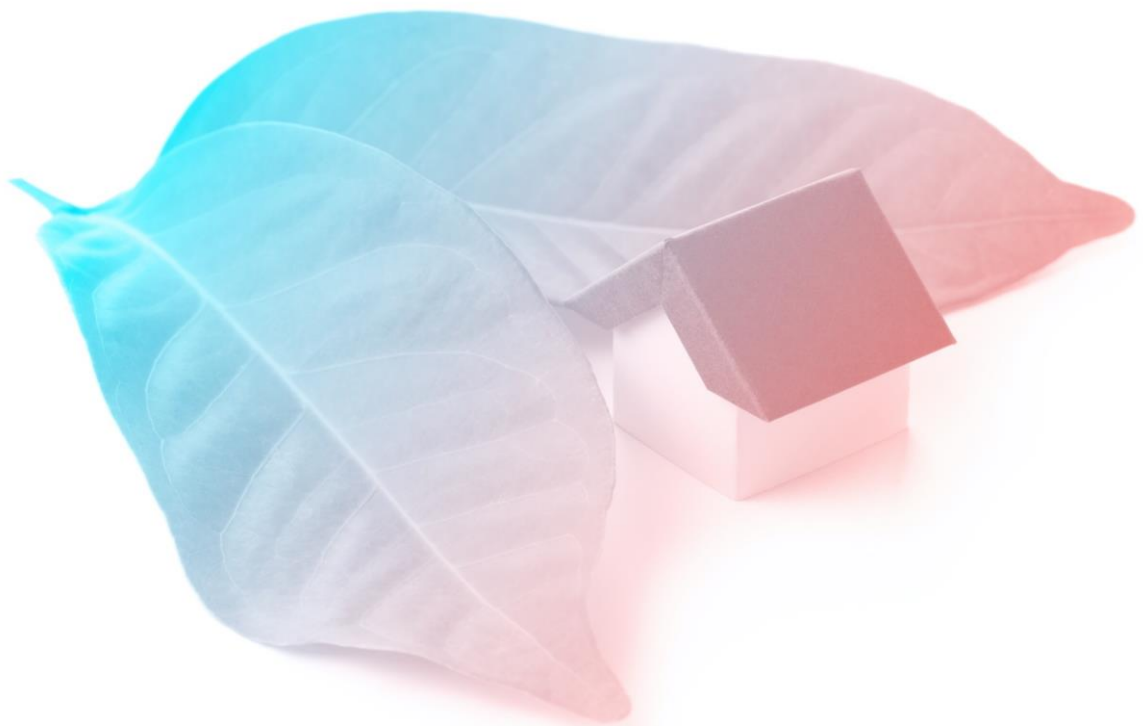




## D1.7 Data Management



Marco Rocchetti (R2M), Giuseppe Longo (R2M), Federica Franzè (R2M)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 869821

## D1.7 Data Management Plan

Summary			
<p>This document contains the first version of the Data Management Plan (DMP) for the Horizon 2020 MINISTOR project (under GA No. 869821). This DMP outlines how data collected, processed and/or generated during the MiniStor project life cycle will be organised, stored and shared.</p> <p>As part of making research data findable, accessible, interoperable and re-usable (FAIR data), the DMP provides information on 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 after the end of the project. Therefore, this plan will ensure that MiniStor data and results are managed suitably, in line with all the requirements set by the European Commission in the "Guidelines on FAIR Data Management in Horizon 2020".</p> <p>Due to the typology of document and management plan, this Deliverable, scheduled for month 18 of the project, is the first version of the Data Management Plan. It will be updated during the project lifespan, according to the evolution of the project. Version two will be produced at M36 and the final version at M54.</p>			
Deliverable Number		Work Package	
D1.7		WP 1	
Lead Beneficiary		Deliverable Author(S)	
R2M		Marco Rocchetti (R2M), Federica Franzè (R2M), Giuseppe Longo (R2M)	
Contributors			
Carlos Ochoa (IERC), Brian Cassidy (CORKCITY), Paraskevi Dimitriadou (DUTH), Driss Stitou (CNRS), Chrysovalantis Kontoulis (CERTH-ITI), Athanasios Nesiadis (CERTH), Robert Gadia (HSLU), Alberto Belda (CARTIF), Adriana Coca (ENDEF), Konrad Zdun (ENETECH), Carmen Coteló (FEUGA), Zoltán Pásztor (WOODSPRING), Maurizio Zaglio (SUNCH), Eszter Hajdu (EMI).			
Planned Delivery Date		Actual Delivery Date	
31/04/2021		15/04/2021	
Type of deliverable	ORDP	Open Research Data Pilot	X
Dissemination Level	CO	Confidential, only for members of the consortium (including the Commission)	
	PU	Public	X

## **DISCLAIMER**

The opinion stated in this deliverable reflects the opinion of the authors and not the opinion of the European Commission.

All intellectual property rights are owned by MINISTOR consortium members and are protected by the applicable laws. Reproduction is not authorised without prior written agreement.

The commercial use of any information contained in this document may require a license from the owner of that information.

## Table of Contents

List of Tables.....	3
List of Images.....	3
List of Acronyms .....	4
1. Introduction .....	5
2 Methodology.....	6
2.1 Data Availability and Open Access .....	6
2.1.1 Classification of available data .....	7
2.1.2 Open Access to publications .....	8
2.1.3 Open data and FAIR principles .....	9
2.1.4 Copyright licences.....	10
2.2 Data storage and sharing.....	10
2.3 DMP template .....	10
3 MiniStor Data Management Plan .....	12
3.1 Data Summary .....	12
3.2 Fair Data.....	15
3.4 Data Security .....	16
3.5 Ethical aspects .....	17
4 Conclusions .....	17
ANNEX I.....	18

## List of Tables

Table 1 - Classification of available data .....	8
Table 2 - FAIR principles <sup>9,10</sup> .....	9
Table 3 - DMP template .....	11
Table 4 Data summary (All the details in Annex I) - CO = Confidential; OA = Open Access .....	13

## List of Images

Figure 1 - Open access to scientific publications and research data .....	7
---	---

## List of Acronyms

Acronym	Meaning
CC	Creative Commons
CO	Confidential
EC	European Commission
EU	European Union
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Re-usable
GA	Grant Agreement
H2020	Horizon 2020
IPR	Intellectual Property Rights
M	Month
NDA	Non-Disclosure Agreement
OA	Open Access
ORD	Open Research Data
PR	Private
RP	Reporting Period
T	Task
WP	Work Package

## 1. Introduction

This document portrays the Data Management Plan (DMP) for the Horizon 2020 project “MINISTOR: Minimal Size Thermal and Electrical Energy Storage System for In-Situ Residential Installation”. The principal aim of this document is to define a strategy for the management of all the data collected, processed and generated during the project, in compliance with the EC guidelines<sup>1,2</sup>, including:

- the handling of research data during and after the end of the project;
- the identification of the data to be collected processed and/or generated, including research data, environmental and economic evaluations;
- the methodology and standards that will be applied;
- whether the data will be shared and made open access, and how;
- how the data will be curated and preserved;
- the management of databases to collect data from the project; and
- the management of internal requirements of the partners.

MINISTOR will follow a proactive strategy for internal data and knowledge management. The DMP is scheduled for M18; however, it is intended to be a living document, evolving throughout the project duration. Thus, updates to the DMP are planned for M36 and M54, which will reflect any important changes to the project that may occur due to inclusion of new datasets, changes in consortium policies or external factors. As established in Article 29<sup>3</sup> of the GA, MINISTOR project falls under the requirements of Open Research Data Pilot (ORDP)<sup>1,4</sup> and will therefore seek to openly disseminate its research results, except in those cases where it goes against the legitimate interests of the project partners (i.e., exploitation and/or protection of results), or there will be a high administrative burden for the disclosure of a dataset or limited worth to other users. The aim of participating in the ORD pilot is to “improve and maximise access to and re-use of research data generated by the project”<sup>1</sup>, taking into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security, as well as data preservation questions. The data resulting from the research and implementation of the project is what will validate the impact of MINISTOR and the conclusions drawn in the scientific publications arising.

Many benefits may come with the sharing of knowledge and data from publicly funded research to the whole research community, the commercial sector and the civil society, and thus it is intended that, where possible, these data will be made available through open access repositories. Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.

---

<sup>1</sup> EC - DG RTD. Guidelines on FAIR Data Management in Horizon 2020, version 3.0, 26 July 2016, in: [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf) accessed on 11/01/2021

<sup>2</sup> EC - DG RTD. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, version 3.2, 21 March 2017, in: [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf) accessed on 11/01/2021

<sup>3</sup> Grant Agreement 869821 - Article 29 Dissemination of results – Open Access – Visibility of EU funding

<sup>4</sup> OpenAIRE. What is the EC Open Research Data Pilot? <https://www.openaire.eu/what-is-the-open-research-data-pilot> accessed on 12/01/2021

## 2 Methodology

According to the "Guidelines on FAIR Data Management in Horizon 2020"<sup>1</sup>, the EC is running a flexible pilot under Horizon 2020 called the Open Research Data Pilot (ORD pilot) with the aim of maximising access to data generated by H2020 projects, and encouraging its re-use. Prior to 2017, the scope of this pilot was limited to selected areas; however, it was reviewed and extended to cover all the thematic areas of H2020 Framework Programme. Open access to research data thereby becomes applicable by default in MINISTOR, although the EC also recognises that there are good reasons to keep some or even all research data generated in a project closed, thus providing opt-out possibilities at any stage. The ORD pilot follows the principle "*as open as possible, as closed as necessary*" and focuses on encouraging sound data management as an essential part of research best practices<sup>5</sup>. MINISTOR, as a grant beneficiary, is thus required to:

- Develop (and keep up-to-date) a DMP which will specify whether the data will be made open access;
- Deposit its data in a research data repository (e.g., Zenodo<sup>6</sup> online OA repository);
- Take measures to enable third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) this research data;
- Provide information about the tools needed to validate the research results.

For this purpose, all the MINISTOR project's partners have been requested to fill a survey including the questions stated in the template provided by the EC guidelines<sup>1</sup> (summarized in Section 2.3). The partner's answers to the survey have served for the development of this first version of the DMP. As the execution of the project progresses and new information and data are generated, the DMP will be updated in the context of the periodic reviews envisaged within the GA (M18, M36, M54).

### 2.1 Data Availability and Open Access

Open access (OA) is defined as the practice of providing on-line, free-of-charge access to *scientific information*, which in the context of research and innovation can refer to peer-reviewed scientific research articles (published in academic journals), or research data (data underlying publications, curated data and/or raw data). As outlined in the "Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020"<sup>2</sup>, broader access to scientific publications and research data serves a number of purposes:

- i) Improve the quality of research by building on a stronger body of existing work.
- ii) Increase efficiency of research by encouraging collaboration and reducing duplication of effort.
- iii) Speed up innovation (faster progress to market) by reducing barriers to information flow.
- iv) Enhance the transparency of scientific progress, easing the involvement of citizens and society.
- v) Allow other researchers, industry or citizens to access and re-use information that has been paid for with public money and hence should not have to be paid for again each time it is accessed.

The EU wants to improve access to scientific information and boost the benefits of public investment in research funded under Horizon 2020. This means making publicly-funded scientific information available online, at no extra cost, to European researchers, innovative industries and the public, while ensuring that it is preserved in the long term<sup>7</sup>. These OA principles are translated into specific requirements in the model GA for H2020 projects. As outlined in *Figure 1*, one of the first decisions to be made in regard to research

<sup>5</sup> EC. H2020 Online Manual - Data Management. [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm) accessed on 11/01/2021

<sup>6</sup> Zenodo - open access repository. <https://zenodo.org/> accessed on 14/01/2021

<sup>7</sup> EC. H2020 Online Manual - Open Access. [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm) accessed on 11/01/2021

dissemination is whether to publish the research findings of the project or to protect some aspects for commercial exploitation. This process creates a series of interdependencies between MINISTOR's technical work program (WP2-WP6), the strategic development roadmap and the communication, dissemination and exploitation work packages (WP7-WP8), and the project management procedures (WP1).

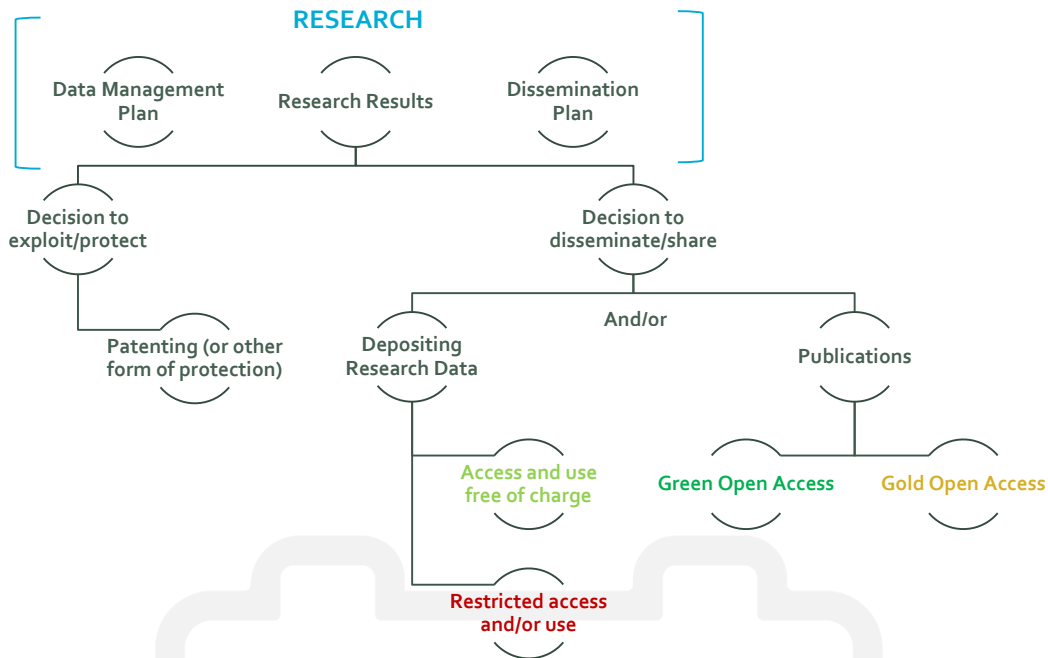


Figure 1 - Open access to scientific publications and research data

For each research result of the project there should be an evaluation whether to disseminate/share or to exploit/protect it. As specified in the GA, “the beneficiaries do not have to ensure open access to specific parts of their research data, if the achievement of the action’s main objective would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access”. The Communication and Dissemination strategy for MINISTOR (D8.2) will also be defined within WP8, in particular T8.1 and T8.2, whilst the exploitation and identification of suitable protection mechanisms will be identified as part of T8.3 and T8.5 (D8.7 IPR & Exploitation Plan).

### 2.1.1 Classification of available data

All available data, datasets and project information are therefore categorised at this stage in one of three ways defined in the Table 1.

The MINISTOR project will seek to openly disseminate its research findings, except in cases where there are defined exploitable outcomes, privacy concerns or there will be a high administrative burden for a dataset to separate open/consortium/private sections leading to limited worth to other users. The two main aspects of this dissemination approach are OA to scientific publications and OA to research data, which are presented in the following sections of this document.

Table 1 - Classification of available data

Data availability	
Open	Data that is shared for re-use or that underpins a scientific publication.
Consortium	Confidential data that is accessible to all partners, but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
Private	Data that is maintained by an individual partner for their own purposes.

### 2.1.2 Open Access to publications

Open access publishing is defined as the free availability of peer-reviewed scientific publications for any user. There are two different paths to follow this approach, the “green” and “gold” OA routes:

- **Green OA** (self-archiving) is the release/deposition of a final peer-reviewed manuscript or article in an online repository before, alongside or after its publication. Repository software usually allows authors to delay access to the article during an 'embargo period', but OA must be ensured within a maximum of 6 months.
- **Gold OA** (open access publishing) is the immediate publication of an article in open access mode through a publisher/journal website. OA must be granted at the latest on the date of publication, and a machine-readable electronic copy must also be deposited in a repository. In the gold route, the payment of publication costs is shifted away from subscribing readers. This model is often based on one-off payments by authors, but these costs (Article Processing Charges, APCs) are eligible for reimbursement during the duration of the project as part of the overall project budget.

Nevertheless, the EC has recently launched the Open Research Europe<sup>8</sup> publishing platform, where H2020-funded researchers can submit scientific papers for peer reviewed publication, fully compliant with all the GA obligations in terms of embargo periods, repositories and provision of OA in a single process, and free of charge. This alternative will be considered by the MINISTOR project's partners.

The requirement to make OA the scientific publications derived from the project do not imply an obligation to publish results. The decision to publish is entirely up to the H2020 grant beneficiaries. Thus, OA becomes an issue only if publication is chosen as a means of dissemination. Moreover, the decision on whether to publish through OA comes after the more general decision on whether to publish directly or to first seek protection, OA hence does not affect the decision to exploit research results commercially.

The MINISTOR consortium will encourage OA publishing of scientific publications. Considering the technological focus and the academic composition of MINISTOR, the consortium partners expect to deliver up to 5 publications (articles) in high-impact journals. Additionally, the project may provide other types of scientific publications (not peer-reviewed), such as monographs, books, conference proceedings and grey literature (informally published written material not controlled by scientific publishers, e.g., reports, press releases). In order to comply with the OA requirements, MINISTOR's dissemination leader (R2M) will establish and maintain a Zenodo<sup>5</sup> online repository for self-archiving of publications and research results, ensuring that they can be found, read, downloaded and printed. The EC guidelines<sup>2</sup> indicate that OA should

<sup>8</sup> EC. Open Research Europe. <https://open-research-europe.ec.europa.eu/> accessed on 19/02/2021



also be provided, through the repository and in a standard format, to the bibliographic metadata<sup>9</sup> that identify the deposited publication. The purpose is to make it easier to find publications and ensure that EU funding is acknowledged. Article 29 of the GA also mentions that H2020 grant beneficiaries (ORD pilot<sup>3</sup> participants) must aim to deposit, at the same time as the publication, the research data needed to validate the results presented in the deposited scientific publications ('underlying data'), other curated and/or raw data not directly attributable to a publication, including the associated metadata. This requirement is not related to the openness of the data but to data management.

### 2.1.3 Open data and FAIR principles

Open access to research data refers to the right to access and reuse digital research data under the terms and conditions set out in the GA; research data consisting of information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation.

"Open data" is normally defined as openly accessible research data that the user can access, mine, exploit, reproduce and disseminate free of charge. As stated in the EC guidelines<sup>1,4</sup>, all the data that MINISTOR will make publicly available (open) will comply with the FAIR principles<sup>10,11</sup> (Table 2) in order to ensure sound data management, as an essential part of research best practice.

Table 2 - FAIR principles<sup>9,10</sup>

<b>F</b>	<b>Findable</b>	Discoverable with metadata, identifiable and locatable by means of a standard identification mechanism.
<b>A</b>	<b>Accessible</b>	Always available and obtainable; even if the data is restricted, the metadata is open.
<b>I</b>	<b>Interoperable</b>	Both syntactically parse able and semantically understandable, allowing data exchange and reuse between researchers, institutions, organisations or countries.
<b>R</b>	<b>Re-usable</b>	Sufficiently described and shared with the least restrictive licences, allowing the widest reuse possible and the least cumbersome integration with other data sources.

FAIR is a set of principles, not a standard. Rather, the EC supports FAIR data as a framework to follow when designing a DMP. It should be noted that participating in the ORD pilot<sup>3</sup> and following the FAIR principles does not mean opening up all the project's data. Data can be FAIR but not open (e.g., data could meet the FAIR principles, but be private or only shared under certain restrictions), whilst open data may not be FAIR

<sup>9</sup> Metadata is "data that provides information about other data". In other words, it is "data about data". Metadata is used to summarize basic information about data which can make tracking and working with specific instances of data easier.

<sup>10</sup> Force 11. The FAIR Data Principles, in: <https://www.force11.org/group/fairgroup/fairprinciples> accessed on 08/01/2021

<sup>11</sup> Wilkinson, M. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18> accessed on 08/01/2021

(e.g., publicly available data may lack sufficient documentation to meet the FAIR principles, such as licensing for clear reuse)<sup>12</sup>.

Datasets for dissemination of the MINISTOR project will be open access by default, at the very least to validate scientific publications. However, not all the project work packages will produce datasets that are intended for public dissemination; selected parts of the data created and stored during the project are to be kept confidential.

#### 2.1.4 Copyright licences

When material is widely shared, copyright licences protect the authors of the work and grant specific rights to publishers and others to use this work. The EC encourages authors to retain their copyright and grant adequate licences to publishers. Creative Commons (CC) provides legal tools and licensing solutions to enable OA in these circumstances, with CC-BY and CC0 enabling re-use by third parties.

When a project's research findings (i.e., facts, data) are published in a journal or other scientific outlet, there should be consideration of the copyright agreement with the publishers, which may involve an embargo period. Submission to the chosen repository requires the authors to agree to a non-exclusive distribution licence, and a Creative Commons licence may be added at this stage. This type of licence is a good legal tool for providing open access in its broadest sense.

## 2.2 Data storage and sharing

Data will be stored and shared based on the different categorisation defined as private, consortium, open:

- **Private data** will be stored locally by the individual owners on their own servers. This type of data storage and sharing process is not in the scope of MINISTOR DMP; however, each partner will ensure the implementation and compliance with their internal data protection processes/policies and national/European regulations.
- **Consortium (confidential) data** will be stored and shared through the project's internal archive, which has been set up using the cloud storage space provided by the project coordinator in Microsoft Teams, as stated in D1.1 "Governance Structures and Methods". This online repository is used for storing and sharing all project related documents/information via the private server, allowing all the project partners to download, archive and exchange project-related data during the whole project execution.
- **Open data** will be made available through Zenodo open data online repository, to be established and maintained by MINISTOR's partner R2M. This storage will ensure that the publicly available (open) scientific publications and research results from the MINISTOR project can be searched, read online, distributed, downloaded and printed.

## 2.3 DMP template

The following template (Table 3) provides a summary of the DMP issues to be addressed, as outlined in the EC guidelines<sup>1</sup>, which will be applied to MINISTOR's datasets as they are developed across the project's lifetime.

---

<sup>12</sup> OpenAIRE. How to make your data FAIR. <https://www.openaire.eu/how-to-make-your-data-fair> accessed on 12/01/2021

Table 3 - DMP template

DMP component	Issues to be addressed
<b>1. Data summary</b>	<ul style="list-style-type: none"> <li>• State the purpose of the data collection/generation</li> <li>• Explain the relation to the objectives of the project</li> <li>• Specify the types and formats of data generated/collected</li> <li>• Specify if existing data is being re-used (if any)</li> <li>• Specify the origin of the data</li> <li>• State the expected size of the data (if known)</li> <li>• Outline the data utility: to whom will it be useful</li> </ul>
<b>2. FAIR Data</b> <b>2.1. Making data findable, including provisions for metadata</b>	<ul style="list-style-type: none"> <li>• Outline the discoverability of data (metadata provision)</li> <li>• Outline the identifiability of data and refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)<sup>13</sup>?</li> <li>• Outline naming conventions used</li> <li>• Outline the approach towards search keyword</li> <li>• Outline the approach for clear versioning</li> <li>• Specify standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how</li> </ul>
<b>2.2 Making data openly accessible</b>	<ul style="list-style-type: none"> <li>• Specify which data will be made openly available? If some data is kept closed provide rationale for doing so</li> <li>• Specify how the data will be made available</li> <li>• Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g., in open-source code)?</li> <li>• Specify where the data and associated metadata, documentation and code are deposited</li> <li>• Specify how access will be provided in case there are any restrictions</li> </ul>
<b>2.3. Making data interoperable</b>	<ul style="list-style-type: none"> <li>• Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.</li> <li>• Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?</li> </ul>
<b>2.4. Increase data re-use (through clarifying licences)</b>	<ul style="list-style-type: none"> <li>• Specify how the data will be licenced to permit the widest reuse possible</li> <li>• Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed</li> <li>• Specify whether the data produced and/or used in the project is usable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why</li> </ul>

<sup>13</sup> A DOI is a persistent identifier for a document that can be handled by a resolution service to direct communications to the correct server. Developed by the International DOI Foundation ([www.doi.org](http://www.doi.org)).

	<ul style="list-style-type: none"> <li>• Describe data quality assurance processes</li> <li>• Specify the length of time for which the data will remain re-usable</li> </ul>
<b>3. Allocation of resources</b>	<ul style="list-style-type: none"> <li>• Estimate the costs for making your data FAIR. Describe how you intend to cover these costs</li> <li>• Clearly identify responsibilities for data management in your project</li> <li>• Describe costs and potential value of long-term preservation</li> </ul>
<b>4. Data security</b>	<ul style="list-style-type: none"> <li>• Address data recovery as well as secure storage and transfer of sensitive data</li> </ul>
<b>5. Ethical aspects</b>	<ul style="list-style-type: none"> <li>• To be covered in the context of the ethics review, ethics section of Description of the Action and ethics deliverables. Include references and related technical aspects if not covered by the former</li> </ul>
<b>6. Other</b>	<ul style="list-style-type: none"> <li>• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)</li> </ul>

### 3 MiniStor Data Management Plan

This section presents the baseline for MINISTOR's DMP, outlining how the data collected, processed and/or generated during the project life cycle will be organised, stored and shared following the FAIR principles. The aim of this plan is to ensure that MINISTOR's information, data and results will be managed suitably. To this end, the project partners have filled a survey in answer to the questions stated in the template provided by the EC in the "Guidelines on FAIR Data Management in Horizon 2020"<sup>14</sup> (summarized in Table 3). The answers to the survey have been collected and analysed by R2M and are compiled and described in the following sections of this document. In this first version of the DMP (M18), not all the template's questions have been answered in detail and the answers are not totally harmonized. The DMP is intended to be a living document, in which information and additional data will be made available on a finer level of granularity through updates as the implementation of the MINISTOR project progresses, and when significant changes occur. This plan will be updated in the context of the periodic reviews envisaged within the GA (M18, M36, M54).

#### 3.1 Data Summary

MINISTOR's project information ("I") and datasets ("D") are reported in a summarized form below (in Table 4). This initial list of datasets has been identified based on the data collected during the first 18 months of the project and the datasets identified as part of the proposal and project planning. At the time of this deliverable submission, 11 technical datasets and 5 project information datasets are foreseen. Although the MINISTOR consortium is committed to OA and public dissemination, not all identified datasets have been classified as OA. There are 11 datasets classified as Consortium (CO), to be kept confidential among the project partners and the EC, although part of these datasets may be made OA at some point (e.g., for validation of the results presented in the public deliverables).

<sup>14</sup> EC - DG RTD. Guidelines on FAIR Data Management in Horizon 2020, version 3.0, 26 July 2016, in: [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf) accessed on 11/01/2021

Table 4 Data summary (All the details in Annex I) - CO = Confidential; OA = Open Access

No.	Dataset description	WP/Task	Class	Data manager
I1	Administrative information not for dissemination purposes: Periodic reports, confidential contact information, meeting minutes, meeting presentations and similar items.	WP1	CO	IERC
I2	Dissemination materials: public deliverables and scientific publications.	WP1, WP8	OA	IERC, FEUGA
I3	Communication materials: newsletters, brochures, posters, videos, and similar items	WP8	OA	FEUGA
I4	Personal Data <sup>15</sup> : of any single person or group of people, either within (i.e., project participants) or external to the Consortium (i.e., stakeholders database), who have in some way actively provided MINISTOR with useful information that could potentially be of some value.	WP1 WP7 WP8	CO	IERC, R2M, FEUGA
I5	Personal Data <sup>14</sup> : of external persons and institutions interested in passively receiving information from or about MINISTOR, such as the periodic newsletter or specific queries received through the website contact form.	WP8	CO	FEUGA
D1	Processed data related to energy modelling in formats of .csv (text) and from energy model vendor format. The processed data report analysis of energy trends and performance of different alternatives.	WP3 T3.6	OA	IERC, all
D2	Lists of demo site sensors measuring internal and external environment and their IP addresses.	WP6 T6.1	CO	HSLU Pilots
D3	Data relative to TCM unit in particular the data relative to the reactive material assessment, the operating procedure, control/command of the TCM unit and other data related to the TCM technology developed in MiniStor	WP3 WP4	CO	HSLU CNRS SUNCH CNRS
D4	Demo site device data/metadata. Weather data, performance output, numeric and date data from sensors/devices and forecasting module. Textual data for demo site and user information (Json format for data dissemination through systems). Safety procedure at pilot site	WP5 WP6 T6.3 WP7	CO	CERTH, WOOD, DUTH, USC, Cork
D5	Draft form of reports, presentations, diagrams and office documents created as well as presentations for the purpose of demonstrating work and explaining concepts of MiniStor project.	WP5 WP6 WP7	CO	CERTH, WOOD, DUTH, USC, Cork CARTIF R2M

<sup>15</sup> The treatment of personal data in MINISTOR will be in compliance with the EU General Data Protection Regulation (GDPR 2016/679 <https://eur-lex.europa.eu/eli/reg/2016/679>).

D6	Raw data generated by models and algorithms used in TRANSYS, MatLAB, ASPEN or Simulink used to perform simulations and calculation, understand & define system operation & characteristics.	WP2 T2.1, 2.4 WP3 T3.1, T3.5 WP6 T6.3, T6.6	CO	CERTH CARTIF
D7	Data reporting thermodynamic calculations and hydraulic drawings	WP2 T2,5 WP3 T3.2 WP4 T4.2 WP6 T6.1, T6.5	CO	HSLU
D8	Operation and control diagrams, Low level and high level control algorithms, electrical diagrams of actuators and control devices, SHEMS architecture. PLC's executable files	WP3 T3.1, T3,5 WP5 T5.1, T5.2, T5.3, T5.4	CO	CARTIF
D9	Raw experimental data from PVT prototypes manufactured and tested in EndeF facilities. Performance's data set	WP3 T3.3, T3.4	CO	EndeF
D10	Market analysis and statistic data used to evaluate market potential, business plan as well as business modelling. List of data related to the project results.	WP7	CO	R2M
D11	List of standards and protocols	WP2 T2.3, T2.5 WP6 T6.2, T6.3	CO	EMI

## 3.2 Fair Data

This section describes how MINISTOR research outputs will be organised so they can be more easily accessed, understood, exchanged and reused. The EC promotes FAIR data to maximise the integrity and impact of their research investment. However, not all the project's data can be open, as there are protection, confidentiality and security obligations which apply, as reflected in the GA and Consortium Agreement. Nevertheless, the MINISTOR consortium will follow the ORD pilot's approach described as "as open as possible, as closed as necessary", having in mind that providing other researchers with access to your research data facilitates knowledge discovery and improves research transparency.

### 3.2.1 Making data findable, including provisions for metadata

MINISTOR will aim to make its research data findable by ensuring it:

- has a **persistent identifier** (PIDs), which allow to unambiguously identify data and facilitates data citation. An example of a PID is a Digital Object Identifier (DOI). Zenodo repository has been selected (proposed at this stage) for depositing MINISTOR data because it assigns PIDs.
- has **rich metadata**, which supports findability, citation and reuse, also providing important context for the interpretation of the data and making automated analysis easier. MINISTOR will aim to follow standard metadata schemes, when possible.
- is **searchable and discoverable online**.

More details about the discoverability measures expected for each of the datasets are presented in the Annex. At this stage of the project, several partners could not specify which data and metadata standards will be used during their research activities, which have not yet started. For the same reason, some data format and metadata can be listed more time. Throughout the project execution, R2M will provide guidance to the rest of the partners on how to make their data findable with the help of data and metadata standards.

### 3.2.2 Making data openly accessible

MINISTOR partners intend to make their data accessible by ensuring it is **retrievable online using standardised protocols**; and has **restrictions in place** if necessary.

As previously mentioned, not all data will be made openly accessible in MINISTOR, but it will be as FAIR as possible, even when restricted. However, if open access is allowed, MINISTOR research data will be retrievable without the need for specialised protocols. Table 4 indicates whether each dataset will be shared and made open access. The Annex shows more detailed information about how the project partners will make their data accessible. R2M will ensure that the project's open data can be accessed online through Zenodo<sup>5</sup> repository.

### 3.2.3 Making data interoperable

In MINISTOR, data will be made interoperable by using common formats and standards, controlled vocabularies, community-agreed schemas, keywords or ontologies, where possible. This will allow the data generated within MINISTOR to be integrated with other data, applications and workflows. When possible, data will not be created with proprietary software, making it available in OA formats.

The answers provided by the partners in regard to the interoperability of their data are gathered in the Annex. The survey specifically requested information about any specific tools (including software) related to the data, and whether these respected data accessibility and interoperability. R2M will evaluate the

formats, standards and tools that will be used by the partners during the execution of their activities in order to provide tailored guidance to the partners and determine if any actions need to be taken (i.e., change of format) to improve the interoperability of the MINISTOR project's data. These actions, if required, will be specified in the upcoming versions of the DMP.

### 3.2.4 Increase data re-use (through clarifying licences)

MINISTOR will aim to make its data reusable by ensuring it:

- is well-documented, by creating documentation (e.g., a README file) to help ensure that the data can be correctly interpreted and analysed by others. For example, such documentation should contain a short description of what data it includes (for each filename); definition of column and row labels, data codes, and measurement units (for tabular data); any data processing steps that may affect interpretation of results; if applicable, a description of what associated datasets are stored elsewhere; or whom to contact with questions.
- has clear licence and provenance information. It is important to have clarifying licenses to govern the terms of the data reuse. The OA guidelines under H2020<sup>10</sup> recommend CC-0 or CC-BY as an effective way to make it possible for others to mine, exploit and reproduce the data.

MINISTOR partners were not able to provide very detailed information on data licensing at this early stage. Nevertheless, it should be noted that large amounts of the data generated within MINISTOR will be consortium-internal data (proprietary license).

## 3.3 Allocation of resources

This section of the DMP addresses the issues related to the costs of making the project's data FAIR and its long-term preservation, and the identification of responsibilities for data management. The partners responsible for each dataset are identified in Table 4. At this stage more partners are indicated as responsible for a dataset. It means that it is not totally defined and it will be finalized in the next DMP versions.

Although R2M will provide guidance on how to follow the FAIR principles (through the DMP), it will be the responsibility of each partner to take measures in order to ensure a sound management of their data. Even though most partners are planning to use their own IT server and network resources for data preservation, all the project data will be also stored and shared (within the consortium) through the project's internal archive (TEAMS), which has been set up by the IERC. No additional costs will be incurred for this purpose. Besides, MINISTOR's publications and open data will be made accessible through Zenodo online OA repository, being the costs related to open access eligible as part of the H2020 grant. The potential value and the duration of MINISTOR's data preservation in the long term is an issue still to be discussed among the partners and to be specified in future updates of the project's DMP. Regarding the costs of making the MINISTOR project's data FAIR, the upcoming updated versions of the DMP will provide more detailed information on whether the project's partners needed to allocate additional resources to comply with the FAIR principles.

## 3.4 Data Security

This section addresses data recovery as well as secure storage and transfer of sensitive data. In order to keep data safe for the long term, MINISTOR has selected a trustworthy repository (Zenodo) that stores the data safely, ensures the data is findable, describes the data appropriately (metadata) and adds license information. Moreover, the partners will store their data in the project's internal archive (TEAMS or Google Drive) as well as in their own servers and data storage systems. The providers of cloud storage and cloud



computation resources and tools employed by the partners (e.g., Google, Microsoft, etc.) guarantee data security, keeping it protected against unauthorized access. The hardware (i.e., laptops) where data is stored are also password protected. Overall, MINISTOR partners will rely on the data security guaranteed by their service providers and their internal security procedures for data storage.

Special attention will be paid to the management of sensitive data throughout the research process. Sensitive data must be protected against unwanted disclosure and its accessibility should hence be safeguarded. The protection of sensitive data may be required for legal or ethical reasons, for issues pertaining to personal privacy, or for proprietary considerations. In particular, MINISTOR will comply with EU regulations regarding personal (GDPR<sup>16</sup>) and sensitive non-personal data (e.g., MINISTOR's confidential data). Sensitive data can still meet the FAIR principles and be processed in a way that the needed protection is guaranteed also in the future.

Sensitive data should be archived under a closed license in a trustworthy repository. However, the metadata could be made public, without making the data itself openly accessible, so it can be findable for other researchers while placing conditions around access to the data. Nevertheless, it should be noted that it is not always necessary to keep all the data, as it will depend on criteria such as its uniqueness, long-term value, reuse potential and the necessity to validate results.

### 3.5 Ethical aspects

The most sensible information identified during the project is the personal data regarding the demo sites' needs of individual users. In this version of the DMP, those aspects will be not deeply considered and will be detailed during the next phases. However, data collection is designed to be anonymized and will only be identified by the geographical location of the demonstration site (e.g. "Cork demo site users"). Only partners responsible for liaising with demonstration sites know the personal details (e.g. names, telephone numbers) of the users. Demo site users have accepted voluntarily to participate in the project after explaining to them the purpose of their participation.

Nevertheless, if any personal data will be considered as part of the data processed within the pilots' activities, and in such a case, each demo site manager will implement GDPR agreements to obtain the consent of the data owners in order to be able to process that data for the purpose of the project, and to share it, only if necessary, with project's partners (for example, recording testimonials regarding their experience with the system).

All the data considered as "business data" should be kept within the consortium partners and used exclusively for the MiniStor project.

All the information obtained from secondary/tertiary sources must be quoted appropriately.

## 4 Conclusions

This document reports a preliminary management tool for the data collected, shared and stored in the MINISTOR project. This DMP and its subsequent updates fulfil the function of ensuring a sound management of the MINISTOR project's data during the project's lifetime and the compliance with the FAIR principles, in accordance with the EC guidelines, the GA and best practices

At the current state of the project 16 datasets are identified. All the data collected in this report comes from the consortium partners and, in this version of the report, they are not totally harmonized by the partners' feedbacks. Therefore, some procedure or management methodology here indicated could be updated in the next version (M36) as well as the inclusion of new databases.

---

<sup>16</sup> GDPR - <https://gdpr-info.eu/>

At month 36 a specific workshop will be organized, with the aim to analyze deeply the final methodology for the storing and sharing of the data at the end of the project. In case of a positive decision, the Zenodo repository will be confirmed and set. These conclusions will be expressed in the final version at M54.

## ANNEX I

In the previous sections, a compilation of the answers has been made in order to provide coherent results among partners. Nevertheless, within the questionnaire, some partners were more detailed with their answers. The following table shows the most detailed answers on the data description question:

IERC Questionnaire	
<b>Data description</b>	<ul style="list-style-type: none"> <li>i) energy modelling data, in formats of .csv (text) and from energy model vendor format</li> <li>ii) analysis of energy trends and performance of different alternatives</li> <li>iii) might be reused for research purposes by further data analysis (statistics)</li> <li>iv) own origin, might be based on data from other partners or from demo site features</li> <li>v) size unknown at the moment, might be hundreds of MBs</li> <li>vi) might be useful to rest of partners to know about system performance</li> </ul>
<b>WPs &amp; Tasks</b>	WP3, Task 3.6
<b>Responsible partner(s)</b>	IERC, all
<b>Dissemination level</b>	OPEN – Data that is shared for re-use or that underpins a scientific publication
<b>Data standards</b>	source code, csv, doc
<b>Standards for any related metadata</b>	They will be findable and discoverable through naming conventions, as in D1.1, search keywords will be determined at the moment of producing the data. Version numbers will be provided.
<b>Specific tools/software (accessibility and interoperability)</b>	csv files can be read by most word processing and notepad programs; there will be accompanying documentation to the software.
<b>Accessibility</b>	Deposition in repository (e.g., Sourceforge)
<b>Deposition of data and associated metadata</b>	Sourceforge and/or Zenodo
<b>With whom is the data shared during the project?</b>	Data will be publicly available unless specified by other project partner there is need to protect certain aspects.
<b>Data licensing</b>	Free of cost license
<b>Preservation after the end of the project?</b>	Yes
<b>How long?</b>	ideally indefinitely, but as programs progress, it might become obsolete (e.g. 10 years)
<b>Who is responsible?</b>	IERC
<b>Where will the data be preserved after the end of project?</b>	Sourceforge or Zenodo

What are the costs of data preservation?	The proposed repositories are free of charge
Ethical and/or legal requirements	Source code is related to code from established programs, but they are open to contributions from external parties
Provisions for data security	Zenodo, Sourceforge and GitHub have data protection measures. However, due to the open-source nature, data needs to be protected against corruption and alteration
CARTIF Questionnaire	
Data description	<ul style="list-style-type: none"> <li>- datasets, data files</li> <li>- documents, spreadsheets, slides, database contents (video, audio, text, images)</li> <li>- models, algorithms, scripts</li> <li>- contents of an application (input, output, logfiles for analysis software, simulation software, schemas)</li> <li>- collections of digital outputs, source code</li> <li>- test responses</li> </ul>
WPs & Tasks	WP5, WP2, WP2, WP6, WP8
Responsible partner(s)	CARTIF, IERC CERTH
Dissemination level	CONSORTIUM - Confidential data that is accessible to all partners, but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
Data standards	No data standards are planned to be used so far.
Standards for any related metadata	No data standards for metadata are planned to be used so far.
Specific tools/software (accessibility and interoperability)	There is a GitHub account associated to the project (not used yet), which is private for now, but could be open access with agreement of all partners, after the project. Excel files can be open using Open Office.
Accessibility	Data meant to be accessible for the Consortium will be included in the Teams platform, where all involved partners have access. Open data will be either included in the website or in public deliverables.
Deposition of data and associated metadata	A GitHub repository was created in the scope of the project but has been barely used.
With whom is the data shared during the project?	No external partners. Data is shared with those partners involved in every activity via provided e-mails or Teams, only.
Data licensing	Data meant to be accessible will be shared in formats that can be accessed using common, open-access software, such as Open Office.
Preservation after the end of the project?	Yes
How long?	At least 10 years
Who is responsible?	The IT department, within CARTIF
Where will the data be preserved after the end of project?	Unknown, but probably in the internal database, which has also a backup in a regional super-computing centre.
What are the costs of data preservation?	Unknown.

<b>Ethical and/or legal requirements</b>	No. No personal data directly included in our activities. All information is anonymised by other partners, if necessary, before being used in our activities.
<b>Provisions for data security</b>	As said, there are periodic backups of our internal database, which is also included in an external database.
<b>CERTH Questionnaire</b>	
<b>Data description</b>	<ul style="list-style-type: none"> <li>i) Demo site, user, and device data/metadata. Numeric and date data from sensors/devices and forecasting module. Textual data for demo site and user information. Json format for data dissemination through systems. Also, Microsoft office documents are being created as well as presentations for the purpose of demonstrating work and explaining concepts.</li> <li>ii) The data collection will be used for the whole system optimization module as well as for visualization and evaluation of components and of the whole system performance (forecasting module, KPIs). Also, documents are created for demonstration and explanation purposes.</li> <li>iii) We will re-use data gathered from MiniStor devices that were installed on demo sites prior to the MiniStor's system.</li> <li>iv) The devices installed per each demo site individually and information from demo sites' owners. Also, MiniStor's partners involved in our tasks.</li> <li>v) The data size depends on the data collection time. As time progresses, more and more data will be collected, thus utilizing more saving space. Generally, recorded data would be in the spectrum of one to two dozen of kilobytes and they will be recorded every minute initially to test system's endurance. Then, this period will increase as more data are collected.</li> <li>vi) The data could be utilized from the site's residents (general audience) as well as the more technically inclined audiences to get system insights. Files generated for demonstration purposes could be used from an audience that wants to replicate the system.</li> <li>vii) Types of data: datasets, data files, documents, spreadsheets, slides, models, algorithms, interviews, questionnaires, dissemination material. Formats of data: doc, xls, ppt, pdf, dat, csv, png, jpeg, m (MatLab), bkp &amp; apwz (ASPEN), slx (Simulink), dwg, vsd.</li> <li>viii) Purpose: implement activities related to the project tasks (perform calculations &amp; simulations, understand &amp; define system operation &amp; characteristics, present results, define demo site characteristics / project risks / users evaluation of MiniStor functionalities / partners' needs &amp; requirements).</li> <li>ix) Many data are used in the framework of several project tasks. Models &amp; algorithms produced by the current entity, may be used for further future development and incorporation of their results in dissemination activities.</li> <li>x) Partners within MiniStor.</li> <li>xi) The majority are up to several MB. Some large datasets are up to 500MB. vi) To project partners.</li> </ul>
<b>WPs &amp; Tasks</b>	T1.2, T1.4, T2.1, T2.2, T2.4, T3.1, T3.2, T3.3, T3.4, T3.5, T3.6, T4.1, T4.2, T4.3, T4.4, T4.5, T5.1, T5.2, T5.3, T5.4, T6.1, T6.2, T6.3, T6.5, T6.6, T7.1, T7.2, T7.3, T7.4, T7.5, T7.6, T7.7, T8.1, T8.2
<b>Responsible partner(s)</b>	CERTH as the host of the database that the data are stored and involved demo sites that provide us with the data. The partner list includes CERTH, Woodspring, DUTH, USC, Cork.
<b>Dissemination level</b>	CONSORTIUM - Confidential data that is accessible to all partners but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).

<p><b>Data standards</b></p>	<p>Most of the produced data are interoperable and compliant with open software applications. The main data format that we intend to use are: comma separated values (.csv), JavaScript object notation (.json), and, datasets (MongoDB -NoSQL, InfluxDB - Timeseries). We aim to adhere only to well-known data formats that can be used from any platform. Access to data from the databases will be provided through an API for security reasons using the json protocol. Considering that all the saved data are in standard and well-established formats, the data are interoperable. No standard vocabularies are foreseen to be used. If needed, mapping to commonly used ontologies can be provided (prerequisite: to have the necessary time to do that).</p>
<p><b>Standards for any related metadata</b></p>	<p>No standards for any related metadata are used. No identification mechanisms are utilised (data do not ease findable). In D1.6 (Quality and Risk Management Plan) specific guidelines for the naming of files are given (i.e., the name includes the related WP and / or Task, date, version number etc.). No search keywords will be provided, except from publications. The guidelines of D1.6 regarding version numbers are followed. No metadata are foreseen to be created.</p> <p>Data can be indexed by using the date and time of recording. Also, data can be indexed through specified endpoints of the RESTful API created for communication with the databases.</p> <p>Endpoint naming is based on each individual device and the assigned unique ID. The unique ID is a sequence of numbers and letters that is bound to a single device. Those unique IDs will be described in relevant deliverables. In general, data from the demos and the sensors/devices will be described in task relevant deliverable documents.</p> <p>Code will be maintained with the help of Gitlab and Open Project that provide version control.</p> <p>In addition to the above metadata, the IoT platform will use HTML to present to the user with the collected data and some information explaining their origin (e.g., electricity consumption, demand-response event, etc.). Scripts will be used for the dissemination of the data within the system components. Jupiter notebooks might be incorporated for the forecasting modules.</p>
<p><b>Specific tools/software (accessibility and interoperability)</b></p>	<p>Some data are accessed with freeware tools (pdf - Foxit Reader, dat - Notepad ++). However, most of them are accessed with propriety tools (doc - MS Word, xls &amp; csv - MS Excel, ppt - MS Power Point, vsd - MS Visio, dwg - AutoCAD, bkp &amp; apwz - ASPEN Plus, slx - Simulink, m - MATLAB). Usually, no documentation about the needed software is included. It is not possible to include relevant software.</p> <p>Methods or software tools needed to access the data: GitLab, MongoDB, InfluxDB, Python, Microsoft Word, Microsoft Excel (.xlsx files). Access to all data can be given through the RESTful API.</p> <p>Documentation defining the endpoints of the RESTful API is needed, but not for the other methods as their use is standard.</p> <p>No, it is not intended as of now. May subject to change in the future.</p>
<p><b>Accessibility</b></p>	<p>The data will be made available through the secure use of the RESTful API, which includes a user authorization mechanism. Home owner data and generated data under the devices of his dwelling are protected through the GTPR and cannot be made available as his privacy will be violated. Though, some data can be made available as open data in Zenodo platform after necessary permissions.</p> <p>The majority of the data are made accessible within the project by their uploading in MS Teams. For some sensitive data, regarding building drawings, equipment and energy</p>

	consumption an NDA agreement is first required. Only data included in publications will be openly available, after the permission of other partners.
<b>Deposition of data and associated metadata</b>	As of now, the data can be accessed through the front-end of an IoT platform from the relevant users after user authorization. Data may be released as open-data in Zenodo. Documentation may be included into a relevant deliverable. Gitlab and OpenProject will provide access to the project's code to authorized personnel. Data are uploaded to and shared via MS Teams. In case of datasets of considerable size, Dropbox or Wetransfer links are given.
<b>With whom is the data shared during the project?</b>	As of now, the data can be accessed only internally by partners of the consortium. The verification of the right person accessing the data is done through a RESTful API, which allows for user identification/authorization. All project partners are able to access the data uploaded in MS Teams. Sometimes, some data are exchanged via email and remain within the knowledge of specific Task participants. CERTH cannot ascertain the identity of person that access data in MS Teams. The content of scientific publications is first reviewed by the project partners, which can raise objections in case of incorporation of sensitive / confidential data.
<b>Data licensing</b>	Data licensing is not defined yet. A possible example would be to publish an open-dataset in Zenodo under the "Creative Commons Attribution 4.0 International" license. To re-use such data, it would be required to cite the dataset's resource. Data creators are the resident of the demo sites. When substantial data will be collected from each demo site. That would be maybe in one semester to one year from now. The data may be usable from any audience as part of publicly released datasets. There is no expiring day. There will be in the relevant deliverable. Since many data are originating from other project partners, no re-use is possible without their permission. The produced data are not usable by third parties. No data quality assurance processes are described.
<b>Preservation after the end of the project?</b>	Yes
<b>How long?</b>	At least 4 years
<b>Who is responsible?</b>	CERTH
<b>Where will the data be preserved after the end of project?</b>	The final is not yet defined. As of now, the data are stored in CERTH's machinery and only for internal use (except from data to be used/ included in publications). The data will be still accessible with the previously defined methods (RESTful API), thus will satisfy FAIR principles.
<b>What are the costs of data preservation?</b>	Open datasets will be hosted by Zenodo as well as CERTH's internal repository. CERTH has available storage infrastructure that stores historical data and there is no cost associated with that. Not able to answer this question. In general, CERTH/CPERI maintains a storage infrastructure that is enlarged and maintained from the Institute Funding Sources (Research projects, governmental funds)
<b>Ethical and/or legal requirements</b>	Some sensitive data related to demo site characteristics (drawings, equipment, energy consumption etc.) require the signing of an NDA before shared. The same may applies to data related to design features, materials properties and functionalities of MiniStor system and its subcomponents (i.e., PVTs, PCMs etc.). So far, no personal data are collected. For questionnaires regarding user assessment of MiniStor functionalities and identification of stakeholder requirements all provisions of GDPR are taken into account.

Provisions for data security	The data are stored in CERTH's server, which provides secure access to them. This is achieved with an authentication mechanism and with controlled access that allows connections only from the internal network and to specific trusted addresses that have been defined with access policies.
CNRS France Questionnaire	
Data description	DATA relative to TCM unit in particular the data relative to the reactive material assessment, the operating procedure, control/command of the TCM unit
WPs & Tasks	WP4 and WP3
Responsible partner(s)	
Dissemination level	CONSORTIUM - Confidential data that is accessible to all partners but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
Data standards	doc, pdf, xlsx, csv
Standards for any related metadata	none
Specific tools/software (accessibility and interoperability)	none
Accessibility	teams platform of the consortium
Deposition of data and associated metadata	teams platform of the consortium
With whom is the data shared during the project?	all partners
Data licensing	No
Preservation after the end of the project?	Yes
How long?	Two Years
Who is responsible?	CNRS
Where will the data be preserved after the end of project?	CNRS
What are the costs of data preservation?	please contact CNRS
Ethical and/or legal requirements	No
Provisions for data security	None
Cork City Council IE Questionnaire	
Data description	Sensory data on the internal and external environment of demo site. Also, data relating to gas and electricity consumption.
WPs & Tasks	Task 6.1
Responsible partner(s)	Cork City Council, IERC/Tyndall, HSLU, Woodspring, CERTH

Dissemination level	
Data standards	
Standards for any related metadata	Will be determined by task leader
Specific tools/software (accessibility and interoperability)	Will be determined by task leader
Accessibility	Personal data of occupants cannot be shared without their permission
Deposition of data and associated metadata	CERTH
With whom is the data shared during the project?	All partners in task 6 and other partners as may be required
Data licensing	Do not know
Preservation after the end of the project?	Yes
How long?	As determined by Cork City Council in agreement with the occupiers
Who is responsible?	Cork City Council
Where will the data be preserved after the end of project?	Cork City Council
What are the costs of data preservation?	Cork City Council
Ethical and/or legal requirements	No ethical concerns. Data can only be preserved with occupiers' consent and for as long as they consent to it.
Provisions for data security	During life of project data will be stored by CERTH. After the project, all data will be securely transferred to Cork City Council location for long term preservation and curation.
<b>Democritus University of Thrace (DUTH) Questionnaire</b>	
Data description	i) Monitoring energy consumption patterns and accessing performance outputs of various equipment iii) yes, historical data of production of energy from the existing system, by comparing the respective spreadsheets of performance, of the new and the existing equipment iv) the existing data-logging system v) several MBs vi) to the Consortium, as it will be an index of the performance of the new systems.
WPs & Tasks	WP6 / Task 6.3
Responsible partner(s)	Democritus University of Thrace (DUTH)
Dissemination level	CONSORTIUM - Confidential data that is accessible to all partners but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
Data standards	- Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e., adhering to standards for formats, as much as possible compliant with available (open) software



	<p>applications, and in particular facilitating re-combinations with different datasets from different origins)? Yes</p> <ul style="list-style-type: none"> <li>- What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?</li> </ul> <p>Standard S.I units and consistent spreadsheets</p> <ul style="list-style-type: none"> <li>- Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability? Yes</li> <li>- In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes</li> </ul>
<b>Standards for any related metadata</b>	<ul style="list-style-type: none"> <li>- Are the data produced and/or used in the project FINDABLE/discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g., persistent and unique identifiers such as Digital Object Identifiers)? No</li> <li>- What naming conventions do you follow?</li> </ul> <p>Abbreviations of the data labels acquired</p> <ul style="list-style-type: none"> <li>- Will search keywords be provided that optimize possibilities for RE-USE? N/A</li> <li>- Do you provide clear version numbers? Yes</li> <li>- What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.</li> </ul>
<b>Specific tools/software (accessibility and interoperability)</b>	<p>DEOS OPENView is the database's software, which acts as a control panel for the data gathering and activation of several electromechanical equipment</p> <ul style="list-style-type: none"> <li>- What methods or software tools are needed to access the data? Non only permission by a DUTH partner</li> <li>- Is documentation about the software needed to access the data included? N/A</li> <li>- Is it possible to include the relevant software (e.g., in open source code)? Yes</li> </ul>
<b>Accessibility</b>	Via CERTH's IoT platform
<b>Deposition of data and associated metadata</b>	N/A
<b>With whom is the data shared during the project?</b>	The data gathered will be private. Access is allowed only to DUTH Team and CERTH, who operates the platform
<b>Data licensing</b>	<ul style="list-style-type: none"> <li>- How will the data be licensed to permit the widest RE-USE possible? N/A</li> <li>- When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply. N/A</li> <li>- Are the data produced and/or used in the project useable by third parties, after the end of the project? If the re-use of some data is restricted, explain why. N/A</li> <li>- How long is it intended that the data remains re-usable? N/A</li> <li>- Are data quality assurance processes described? N/A</li> </ul>
<b>Preservation after the end of the project?</b>	Yes
<b>How long?</b>	Up to 3 Years
<b>Who is responsible?</b>	DUTH Team
<b>Where will the data be preserved after the end of project?</b>	The data will be available at DUTH's SQL database upon request.
<b>What are the costs of data preservation?</b>	The data will be stored in a SQL database for up to 3 years and there are no further costs for long term preservation.

<b>Ethical and/or legal requirements</b>	<p>No legal/ethical requirements. A consent form has been signed by the users about the use of their data</p> <ul style="list-style-type: none"> <li>- How will the data be acquired? Via an .exe app installed in the PC where the SQL database is installed (CERTH's input might be more useful)</li> <li>- When and where will they be acquired? Acc. to CERTH's requirements, every minute</li> <li>- How will they be stored? Stored in a secure SQL database</li> <li>- Who will be able to access the data? Only DUTH Team and CERTH's platform operator</li> <li>- Is it personal data? Yes</li> </ul>
<b>Provisions for data security</b>	The data is stored in a SQL database whose IP is protected by the University's firewall
<b>ÉMI Questionnaire</b>	
<b>Data description</b>	<ul style="list-style-type: none"> <li>i) <ul style="list-style-type: none"> <li>- contribution to standards and protocols in document format</li> <li>- test results as spreadsheets, eventually digital outputs</li> <li>- videos, images, articles</li> </ul> </li> <li>ii) <ul style="list-style-type: none"> <li>- contributing to the creation of own shipping and maintenance guidelines and propositions for the prototype</li> <li>- gathering test results for the validation of the MiniStor prototype</li> <li>- publishing videos, images, articles as dissemination activity</li> </ul> </li> <li>iii) <ul style="list-style-type: none"> <li>- studying the existing data from standards and protocols in order to create own guidelines and propositions</li> </ul> </li> <li>iv) <ul style="list-style-type: none"> <li>- existing standards and protocols</li> <li>- own measurements</li> </ul> </li> <li>v.) <ul style="list-style-type: none"> <li>not known yet</li> </ul> </li> <li>vi) <ul style="list-style-type: none"> <li>for all partners</li> </ul> </li> </ul>
<b>WPs &amp; Tasks</b>	T2.3, T2.5, T6.2, T6.4, T8.2, T8.3
<b>Responsible partner(s)</b>	SGS, HSLU, FEUGA, IER (however not sure about the meaning of this question)
<b>Dissemination level</b>	CONSORTIUM - Confidential data that is accessible to all partners but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
<b>Data standards</b>	We use standard vocabularies as documents, spreadsheets, presentations, images, interviews
<b>Standards for any related metadata</b>	not applicable
<b>Specific tools/software (accessibility and interoperability)</b>	using Microsoft Word and Excel, no specific tools or softwares
<b>Accessibility</b>	sharing via e-mail or the project's Microsoft Teams
<b>Deposition of data and associated metadata</b>	N/A
<b>With whom is the data shared during the project?</b>	We provide the data available only within the project partners at first stage
<b>Data licensing</b>	N/A

Preservation after the end of the project?	Yes
How long?	not limited, keeping within ÉMI's filing system
Who is responsible?	project manager
Where will the data be preserved after the end of project?	within ÉMI's filing system electronically and some in printed version too
What are the costs of data preservation?	N/A
Ethical and/or legal requirements	No
Provisions for data security	data will be stored and handled within regular company policy
<b>ENDEF Questionnaire</b>	
Data description	Experimental data from PVT prototypes manufactured and tested in EndeF facilities
WPs & Tasks	Task 3.3, Task 3.4
Responsible partner(s)	EndeF
Dissemination level	PRIVATE - Data that is maintained by an individual partner for their own purposes.
Data standards	With the experimental data the performance curves for the PVT collectors are obtained
Standards for any related metadata	The formats mentioned are not applied to the data
Specific tools/software (accessibility and interoperability)	The experimental data were / will be processed in Matlab
Accessibility	Only the performance curves will be provided to the project.
Deposition of data and associated metadata	This part is not applied to Task 3.3/3.4
With whom is the data shared during the project?	Only the performance curves will be provided to the project.
Data licensing	This part is not applied to Task 3.3/3.4
Preservation after the end of the project?	Yes
How long?	2-5 Years
Who is responsible?	EndeF
Where will the data be preserved after the end of project?	EndeF server

What are the costs of data preservation?	There is not information in relation to it
Ethical and/or legal requirements	This part is not applied to Task 3.3/3.4
Provisions for data security	There is no information in relation to it
<b>Enetech Questionnaire</b>	
Data description	<p>i) The work made by Enetech in the MiniStor project is desk-based work. We are mainly involved in WP 2 and WP7. Types of data collected, generated, and processed are mainly text, charts, maps.</p> <p>ii) Data collected is needed to perform necessary analysis such as market potential</p> <p>iii) Probably yes, data once used can be used in another related task iv) Data comes from the literature, statistics provided by relevant offices v) This kind of data is rather small, up to few MB vi) Firstly it can be useful for other partners, investors, but also final users.</p>
WPs & Tasks	T 2.1, T2.4, WP 7 in the future.
Responsible partner(s)	CERTH
Dissemination level	CONSORTIUM - Confidential data that is accessible to all partners but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
Data standards	In connection to previous answer, mainly used data standards are documents (.doc), spreadsheets (.xls), images (.jpeg)
Standards for any related metadata	Enetech do not provide such types of data
Specific tools/software (accessibility and interoperability)	Software needed to our data is MS Office
Accessibility	Data is put into suitable directory in MS Teams.
Deposition of data and associated metadata	Not applicable
With whom is the data shared during the project?	Data is shared with project partners
Data licensing	Not applicable
Preservation after the end of the project?	Yes
How long?	Two Years
Who is responsible?	Task Leaders
Where will the data be preserved after the end of project?	
What are the costs of data preservation?	

Ethical and/or legal requirements	No
Provisions for data security	Data is stored on computers hard drives, as well as there are backups made periodically
<b>R2M Solution Questionnaire</b>	
Data description	Reports in form of .doc file Data collection related to Exploitable result (excel files) Schemes, tables and graphs Business modelling
WPs & Tasks	WP6, WP7, WP8
Responsible partner(s)	R2M, IERC, ER owners
Dissemination level	Mainly confidential (CO)
Data standards	Doc, excel, jpeg, PDF
Standards for any related metadata	Not defined
Specific tools/software (accessibility and interoperability)	Microsoft Office tools
Accessibility	Data will be available by TEAMS
Deposition of data and associated metadata	N/A
With whom is the data shared during the project?	Mainly with the Pilots and ER owners MiniStor consortium External stakeholder
Data licensing	Not defined yet
Preservation after the end of the project?	Yes, data will be preserved in the R2M GDrive
How long?	Not decided
Who is responsible?	Project Manager of the company
Where will the data be preserved after the end of project?	R2M GDrive
What are the costs of data preservation?	N/A
Ethical and/or legal requirements	Not defined yet
Provisions for data security	Data is stored on computers hard drives, as well as there are backups made periodically
<b>FEUGA Questionnaire</b>	
Data description	i) The dissemination and Communication material refers to the following items: web and social media contents, posters, presentations, images, videos, flyers, infographic,

	<p>newsletter, press releases, blog posts, questionnaires, etc. Scientific and technical papers will be key Dissemination elements but produced/published from other WPs and partners.</p> <p>ii) The purpose of the data collection and processing is the generation of public Communication and Dissemination materials as main tools to communicate project and its results to all target audiences to maximize the project impact.</p> <p>iii) Most of the Communication and Dissemination materials will be new, we don't expect to re-use previous data (expect public images or images provided from the partners, e.g. pilot sites).</p> <p>iv) Most of the Communication and Dissemination materials produced will be result the data gathered from the different MiniStor technical WPs. Scientific papers/presentations/diagrams used for Comm&amp; Diss activities (and website) will be produced under the rest of WPs and shared through Comm&amp;Diss activities and channels. Analytics form the web server will be also gathered for analysis and web optimization.</p> <p>v) The final size will mainly depend on the number of videos, images and presentations produced.</p> <p>It will be useful to the whole consortium in order to disseminate their project results, to other related projects and research community (presentations, scientific and technical papers, etc..) and the wide public in general to know about the project topic (infographics, social media, ...).</p>
<b>WPs &amp; Tasks</b>	Task 8.1
<b>Responsible partner(s)</b>	Not defined
<b>Dissemination level</b>	OPEN - Data that is shared for re-use or that underpins a scientific publication.
<b>Data standards</b>	<p>Yes, Communication and Dissemination materials are produced in formats as much as possible compliant with available software applications:</p> <ul style="list-style-type: none"> <li>- web (html, css, javascript,..)</li> <li>- documents (.doc, .txt)</li> <li>- images (.jpeg, .png, etc.)</li> <li>- presentations (.odp, .ppt, .pdf)</li> <li>- spreadsheets (.ods, .xlsx)</li> <li>- interviews (.txt, .odt, .pdf, .mp3, .mp4)</li> </ul>
<b>Standards for any related metadata</b>	<p>Keywords and descriptions (as HTML meta tags) will be added to the MiniStor web site.</p> <p>Specific hashtags are used for Social Media as keywords</p> <p>* Other metadata for scientific and technical papers/reports (other WPs) will be managed by the publications' authors</p>
<b>Specific tools/software (accessibility and interoperability)</b>	<p>Microsoft Word, Excel and Power Point</p> <p>Word Press (Web site management)</p> <p>Canva (design)</p> <p>Adobe InDesign (to be confirmed)</p>
<b>Accessibility</b>	Public Communication and Dissemination materials will be available in the MiniStor Website.
<b>Deposition of data and associated metadata</b>	At least in the MiniStor website <a href="https://MiniStor.eu/">https://MiniStor.eu/</a>
<b>With whom is the data shared during the project?</b>	Communication and Dissemination materials will be public.
<b>Data licensing</b>	Public domain for the Communication and Dissemination materials. MiniStor project will be creator

	* For scientific publications, the 'golden' open access model will be chosen. Authors/creators of these publications will depend on the topic and WPs (acknowledgement of EU funding for MiniStor project always included following the H2020 rules).
<b>Preservation after the end of the project?</b>	Yes
<b>How long?</b>	At least 3 years after the end of the project
<b>Who is responsible?</b>	Public data will be available on the MiniStor website or/and partners websites/repositories
<b>Where will the data be preserved after the end of project?</b>	The resources for long term preservation have not yet been discussed
<b>What are the costs of data preservation?</b>	The costs for making the data FAIR are those related the hosting and domain of the MiniStor website.
<b>Ethical and/or legal requirements</b>	No ethical or legal issues are foreseen In case of questionnaires dealing with personal data informed consent will be included
<b>Provisions for data security</b>	Communication and Dissemination material is public, so no sensitive data transfer expected. Regular backups
<b>HSLU Questionnaire</b>	
<b>Data description</b>	<ul style="list-style-type: none"> <li>i) Excel, pdf, Visio, PowerPoint, Word, etc.</li> <li>ii) Project management (to organise internally the person months and to do's), Thermodynamic calculations (calculations for approximations, calculate viability of ideas, etc.), hydraulic drawings (in order to know which components are connected), prepare presentations to show results and progress of the project, write deliverables, etc.</li> <li>iii) Probably all the produced data will be re-used. The goal is to review the work, find and write down lessons learned and see different points of view and manners to come to a solution, etc.</li> <li>iv) Most of the data will be produced by ourselves. The data that we will receive is the data coming from the measurements from the demo sites.</li> <li>v) Guess: 100 kB and 100 MB.</li> </ul>
<b>WPs &amp; Tasks</b>	WP2, T2.5, T3.2, T4.2, T6.1, T6.5
<b>Responsible partner(s)</b>	HSLU
<b>Dissemination level</b>	CONSORTIUM - Confidential data that is accessible to all partners but retained within the consortium and subject to the project Non-Disclosure Agreement (NDA).
<b>Data standards</b>	<ul style="list-style-type: none"> <li>i) Yes</li> <li>ii) general standards and specially using legends to explain symbols and data in general</li> <li>iii) Yes</li> <li>iv) Yes</li> </ul>

<b>Standards for any related metadata</b>	i) Not planned, but we will try to do so ii) No standard yet iii) Not planned, but we will try to do so iv) Not planned, but we will try to do so v) Not clear yet
<b>Specific tools/software (accessibility and interoperability)</b>	i) Excel, Spyder, etc. ii) No iii) Yes
<b>Accessibility</b>	Different options: Send by e-mail, uploaded in Teams, uploaded in another tool to transfer bigger datasets
<b>Deposition of data and associated metadata</b>	Not defined yet
<b>With whom is the data shared during the project?</b>	Related partners of each task
<b>Data licensing</b>	i) Not defined yet ii) Not defined yet, but data can be requested iii) Restricted. Now we cannot say if sensible data will be created during the project, so for the moment we are cautious and say restricted. iv) As long as the data es useful for the user v) No
<b>Preservation after the end of the project?</b>	Yes
<b>How long?</b>	If it is useful
<b>Who is responsible?</b>	Project leader of MiniStor in HSLU (internal)
<b>Where will the data be preserved after the end of project?</b>	Internal Servers
<b>What are the costs of data preservation?</b>	To be defined
<b>Ethical and/or legal requirements</b>	i) Depends on the dataset. As written before, most of the data will be created by ourselves and the data of the demo sites will be acquired during the measurement time (installation of measurement devices - M50)
<b>Provisions for data security</b>	To be determined
<b>Sunamp Switzerland GmbH Questionnaire</b>	
<b>Data description</b>	Documents / discussions with Sunamp Ltd and with other project partners
<b>WPs &amp; Tasks</b>	Mainly WP4
<b>Responsible partner(s)</b>	CNRS
<b>Dissemination level</b>	PRIVATE - Data that is maintained by an individual partner for their own purposes.
<b>Data standards</b>	Data ore interoperable (mainly office package); with standard vocabularies.
<b>Standards for any related metadata</b>	No; no convention; it depends on the tools used to open the files and their ability to search for keywords; yes; no metadata



Specific tools/software (accessibility and interoperability)	Office package; no; no
Accessibility	Yes in the teams repository
Deposition of data and associated metadata	Teams' folders
With whom is the data shared during the project?	Internal stakeholders
Data licensing	No use outside project without written permission
Preservation after the end of the project?	Yes
How long?	No idea
Who is responsible?	The creator of the data
Where will the data be preserved after the end of project?	Local laptops and company servers accessible only by internal personnel
What are the costs of data preservation?	By company funding
Ethical and/or legal requirements	Yes, confidentiality
Provisions for data security	In company servers