MULTIPLHY

Multimegawatt high-temperature electrolyser to generate green hydrogen for production of high-quality biofuels

Open Data Management Plan Deliverable D8.6

Grant agreement number 875123
Start: 01/01/2020 – Duration: 60 months

Document Classification

Title	Open Data Management Plan
Deliverable	D8.6
Reporting Period:	RP1 M1-M12
Date of Delivery foreseen	M6
Draft delivery date	18/05/2020
Validation date	30/06/2020

Authors	J. Mougin, CEA
Work package	WP8
Dissemination	PU = public
Nature	R: Document, report
Version	*V1*
Keywords	Data management, accessibility, interoperability, re-usability

Document Validation

Partner	Approval (Signature or e-mail reference)
WP Leader	CEA
Coordinator (PC/SC)	CEA, SF
Others (if applicable)	PW, ENGIE, NESTE

Table of content

Abstract

The Open Research Data Pilot aims to make the research data collected, processed or generated by Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access and respecting Intellectual Property Rules (IPR).
As MULTIPLHY project is part of H2020, all the consortium members have to assume stakeholder role in "Open access to scientific publications" policy.

The MULTIPLHY project data management plan is a "living" document presenting, at each moment of the project, how data collected or generated will be handled during and after the project.

The Data Management Plan (DMP) will include:
- Accounting for IPR, commercial and security issues, the definition of the MULTIPLHY project data that will be shared/made open access.
- The definition of the repository to handle the data during and after the end of the project.

A description of the methodology, including standards, used to curate and preserve data collected, processed or generated.

1. Introduction

Sound data management is a part of research best practice that the European Commission has decided to promote by mean of Open Research Data (ORD) Pilot project. ORD pilot aims to improve and maximize access to and re-use of research data collected, generated or processed by all Horizon 2020 projects, including MULTIPLHY.

The present document, deliverable D8.6 Open Data Management Plan (DMP), is a part of MULTIPLHY WP8: Dissemination, Communication and Exploitation. MULTIPLHY project started in January 2020 and consequently participates in ORD Pilot that means the project has to provide a first version of DMP six month after its start.

The DMP is not a fixed document, but it is likely to evolve during the whole lifespan of the project, serving as a working document. This document will be updated as often as needed during the Project General Assemblies.

The purpose of the current deliverable is to present the first version of the Data Management Plan of the MULTIPLHY project. The deliverable has been compiled with the collaborative work among the coordinator and the consortium partners who are involved in data collection, production and processing. It includes detailed descriptions of all datasets that will be collected, processed or generated in all Work packages during the course of the 60 months of MULTIPLHY project. The deliverable is submitted through the latest guidelines: The Open Research Data Pilot (ORD Pilot). For the methodological part, the latest EC guidelines[1] have been adopted for the current deliverable.

The deliverable is structured into the following sections:

1. An introduction to the deliverable and a brief description on how Data Management is approached in Horizon 2020 (H2020) program along with the importance of it.
2. A description of the methodology used, an analysis of the chapters of the provided template and last the methodological steps followed in MULTIPLHY.
3. A description of the datasets to be used in MULTIPLHY.
4. A DMP revision timetable
5. A summary table with all the datasets included in first MULTIPLHY DMP.

2. Data management in H2020 programs

First of all, it seems necessary to remind that a fundamental distinction has to be made, in the frame of H2020 programs, between peer-reviewed publications, for which open access is an obligation, and open research data, that can be open or closed.

For open research data the commission is running a flexible pilot: Open Research Data pilot that helps H2020 beneficiaries to make their research data Findable, Accessible, Interoperable and Re-usable (FAIR), to ensure it is soundly managed.

Initially limited to some research areas, the Open Research Data pilot has been extended since 2017 to cover all the thematic areas of Horizon 2020 (H2020). Consequently, it will be applied to MULTIPLHY project started in 2020.

The aim of ORD pilot is to improve and maximize access to and re-use of research data collected/ generated/ processed in the frame of H2020 projects. Nevertheless, the goal is not to necessarily open all the research data, but rather to follow the principle "as open as possible, as closed as

---

[1] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

necessary". ORD pilot deals with the need to balance openness and protection of scientific information, Intellectual Property Rights (IPR), privacy concerns, security, as well as data management and preservation issues.

A Data Management Plan (DMP) is a key point to reach the goal of producing FAIR data. A DMP describes the data management life cycle for the data to be collected, generated or processed and should include at least information about: (i) how handle research data during and after the end of the project, (ii) what data will be collected, processed, generated, (iii) what data will be opened and what will remain closed and for what reason, (iv) which methodology and standards will be applied, (v) how data will be curated and preserved (including after the end of the project).

Notice that ORD pilot applies primarily to the data needed to validate the results presented in scientific publications. Other data can be provided on a voluntary basis as stated in the DMP.

The DMP remain overall a 'living" document that has to be kept up to date as the project is running.

A periodical review of the DMP is necessary and has to be defined and discussed between consortium partners.

3. Methodology

3.1. DMP Template

In order to assist the beneficiaries with the completion of the DMP, the EC produced and provided a template that acts as a basis for data description. The template contains a set of questions that beneficiaries should answer with a level of detail appropriate to the project. If no related information is available for a given dataset, then the phrase "*Non-applicable*" or N/A will be used. In the following paragraphs, the main sections and proposed contents of the template are listed and presented, along with the way MULTIPLHY reflects these sections. Finally the table provided in appendix 1 summarizes main issues expressed in the template.

3.1.1 Data summary

In this section, beneficiaries are asked to describe (a) the purpose of the data collection or generation and how this purpose reflects the objectives set in the project as a whole, (b) the types and formats of data that will be generated or collected, (c) the origin of the data, (d) the expected size of the data, and also (e) whether existing data will be re-used and (f) the usefulness of the described datasets.

3.1.2 FAIR data

a) Making data findable, including provisions for metadata

This section includes a description of metadata and related standards, the naming and keywords to be used. In the context of MULTIPLHY the following naming convention will be used for all the datasets of the project. First the project name: MULTIPLHY, then the work package number will be placed, then the date using YYYYMMDD (format that insures direct correspondence between chronological and alphabetical orders), then the ID of the partner who produced the data, then the serial number of the dataset within this work package and last the dataset title, all separated with underscore (MULTIPLHY_Data_<WPno>_YYYYMMDD_< data producer ID >_<serial number of dataset (2 digits)>_<dataset title>). Notice that serial numbering will be performed by 2 digits to make consistent alphabetical and chronological filing

An example can be the following: *MULTIPLHY_Data_WP2_20201222_CEA_01_ stack testing _data*.

However, it has to be noted that this naming convention describes only the general dataset that can contain files of different size and format. The naming of each separate file follows a different naming convention that is proposed by the partner who creates the files.

The use of a standard identification mechanism for the datasets of MULTIPLHY will be decided by the project consortium. If it turns out to be necessary, the use of the Guidelines and standards provided by the International DOI Foundation (IDF) and the DOI system and ISO 26324[2] will be considered.

b)   Making data openly accessible

Which data produced or used in the project has to be made openly available by default? That is the main issue of the section. It also explains why some datasets cannot be made open due to possible, legal, contractual or ethical issues. Some beneficiaries may have decided to keep their data closed. A description of the potential data repositories is also included along with the potential software tools required to access the data.

In the context of MULTIPLHY, the following options for open repositories of data, metadata, documentation or code will be considered: (a) The Registry of Research Data Repositories[3], (b) Zenodo[4], (c) OpenAIRE[5],

In the context of the MULTIPLHY DMP, no arrangement has been made with an identified repository. This will be discussed by the consortium during the upcoming meeting.

c)   Making data interoperable

In this section, data interoperability is detailed for every dataset of MULTIPLHY. Issues such as the allowing of data exchange between researchers, institutions or even countries are covered along with all the technicalities including standards for formats or metadata vocabularies.

The issue of interoperability will be discussed among the consortium members in the next project meeting.

d)   Increase data re-use (through clarifying licenses)

This section describes the licenses, if any, under which data will be re-used in MULTIPLHY. It includes provisions regarding the period when data will be available for re-use and if third parties will have the option to use the data and when.

e)   Allocation of resources

FAIR data management in MULTIPLHY project is led by the project coordinator CEA under WP7 – Communication, Dissemination and Exploitation of project results. All costs related to FAIR data management that will occur during project implementation will be covered by the project budget. Any other cost that may relate to long term data preservation will be discussed among consortium members.

f)   Data security

Data security is of major importance in the MULTIPLHY project. Special attention will be given to the security of sensitive data. The protection of data will be ensured through procedures and appropriate technologies. If data will be kept in a certified repository, then the security standards of that repository will apply.

g)   Ethical aspects

---

[2] https://www.doi.org

[3] http://www.re3data.org/

[4] https://zenodo.org/

[5] https://www.openaire.eu/

With respect to the H2020 ethics self-assessment, the MULTIPLHY proposal and the use case scenarios to be defined will not be concerned with any ethical issue.

h)   Other issues

In this section, other issues not included above can be covered such as the use of other national/funder/sectorial/departmental procedures for data management.

3.2.   Methodological steps in MULTIPLHY

For the 1st version of MULTIPLHY DMP, the following methodological steps were followed:

1. The coordinator, responsible for the implementation of DMP in the frame of WP8 – Communication, Dissemination and Exploitation of project results - sent to all partners an email notifying them about the upcoming deliverable. Contribution was asked from all partners that were involved in any data collection in each task of the WPs. They were asked to answer a questionnaire on which data they were expecting to produce and collect during the project.
2. In parallel, the latest guidelines from the EC regarding data management were sent to all partners to be informed.
3. The project team collaborated efficiently and contributed with the needed information.

The first version of the MULTIPLHY DMP is intended to provide an initial screening of the data to be collected, processed and produced within MULTIPLHY. It is also the first attempt to collect the vision and input from all the partners involved in any data management option. It can be underlined that not all data generated in MULTIPLHY will be public for a matter of confidentiality and industrial and intellectual property. A selection of data which can be made available to the community without affecting confidentiality for partners is provided in below sections. During the upcoming Project Meetings special attention will be given to data management in order to provide further clarifications and conclusions on data management.

4. DMP review timetable

DMP is a living document that have to be adjusted all along the project lifetime. Plenary meeting periodicity seems to be adapted to update the data management plan and ensure its progress to its consolidation.

M6	D8.6 Open Data Management Plan v1
M30	D8.6 Open Data Management Plan v2 (shared between MULTIPLHY partners),
M60	D8.7 Consolidated Data Management Plan

At the end of the Project a consolidated Data Management Plan (D8.7) will be provided that will detail deeply the MULTIPLHY project datasets and the way to access to the data.
This final document will be fully open to public access.

5. MULTIPLHY datasets

Following tables, one for each work package that uses, generates or processes data, address the data management plan issues detailed in present document appendix 2.

5.1. Work package 2: Stack tests at 10 kW in laboratory

DMP component	MULTIPLHY_data_WP2
1. Data summary	Purpose: test and benchmark 3 different stack technologies following a harmonized testing protocol, evaluate their long-term durability, improve the knowledge of degradation mechanisms Data format: xlsx, mat or csv; docx; jpg, tiff, png Re-use of existing data and if yes, how? Yes. SOCTESQA and JRC based testing protocols; Data available at testing partners for comparison Origin of data: Performance and durability tests performed in the frame of MULTIPLHY, according to protocols defined and obtained on the test stations at different partners laboratories Expected size of data: Approx. several Gb To whom might it be useful? Scientific community, private industrial actors
2. FAIR Data	
2.1 Making data findable, including provisions for metadata	Description of the data: Naming will be WP2__YYYYMMDD_<Producer ID>_xx_test results_data
2.2 Making data openly accessible	Harmonized test protocols will be public, as public deliverable D2.1. Results of long term durability tests will be made public, in agreement with what will be reported in public deliverable D2.3. Short term release tests performed at the start of the WP won't be public, but shared between partners as required to ensure optimal progress in the project What method or software tools are needed to access the data? Data will be exported as binary data. It will be possible to export part of them in Excel format to be published as open data. How will the identity of the person accessing the data be ascertained? Information about company, activity will be requested during first access.
2.3 Making data interoperable	Public data of the project for this thematic are interoperable by nature.

DMP component	
2.4. Increase data re-use (through clarifying licenses)	N/A
3. Allocation of resources	All costs related to the data collection and processing are covered by the project budget under WP8
4. Data security	Sensible and non-public data will be stored on a SharePoint available only to the consortium members
5. Ethical aspects	No ethical aspects for MULTIPLHY
6. Other	N/A

5.2.   Work package 3: System design and manufacturing

DMP component	MULTIPLHY_data_WP3
1 Data summary	Purpose: define, engineer and manufacture the HTE system and its auxiliary unit (especially hydrogen processing unit), and define the operation concept in the refinery. Data format: docx, pdf, xlsx, pptx, m, mat, fig, jpeg, … Re-use of existing data and if yes, how? Yes. The MULTIPLHY system belongs to a SF product line strategy. Parts of the documentation will be generic and adapted to the MULTIPLHY system. Documentation related to HTE systems and HPU units already developed by SF and PW will be used as a background. Origin of data: Specification, design report, layouts, technical documentation, factory acceptance test. Expected size of data: GBs To whom might it be useful? Scientific community, private industrial actors, potential customers
2 FAIR Data	
2.1 Making data findable, including provisions for metadata	N/A
2.2 Making data openly accessible	Documents generated in WP3, mainly Specifications, Design, engineering, production and FAT will not be public. The documentation will be available within the consortium. The deliverables of WP3 are confidential and will only be shared with the MultiPLHY consortium members.
2.3 Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licenses)	N/A
3 Allocation of resources	N/A

4	Data security	Sensible and non-public data will be stored on a SharePoint available only for the consortium members
5	Ethical aspects	No ethical aspects for MULTIPLHY
6	Other	N/A

5.3. Work package 4: Integration in refining process

DMP component	MULTIPLHY_data_WP4
1 Data summary	Purpose: Engineering and preparation of installation site, site integration, system commissioning and SAT, and validation of operation procedures Data format: docx, pdf, xlsx, pptx, m, mat, fig… Re-use of existing data and if yes, how? Yes. The MULTIPLHY system belongs to a SF product line strategy. Documentation related to the operation of HTE systems and HPU units already developed by SF and PW will be used as a background. Origin of data: System installation and permissions. Site acceptance test phase Expected size of data: some GB To whom might it be useful? Scientific community, private industrial actors, potential customers
2 FAIR Data	
2.1 Making data findable, including provisions for metadata	N/A
2.2 Making data openly accessible	Documents generated in WP4, related to Site and HTE/HPU commissioning documents, SAT documents and operation procedures won't be public. The documentation will be available within the consortium. All deliverables of WP4 are confidential and will only be shared with the MultiPLHY consortium members.
2.3 Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licences)	N/A
3. Allocation of resources	N/A
4.Data security	Sensible and non-public data will be stored on a SharePoint available only for the consortium members (i.e. all deliverables of WP4)
5. Ethical aspects	No ethical aspects for MULTIPLHY

6. Other	N/A

5.4. Work package 5: Technology validation and demonstration

DMP component	MULTIPLHY_data_WP5
1 Data summary	Purpose: system performance assessment and long term demonstration and degradation evaluation. Data format: docx, xlsx Re-use of existing data and if yes, how? No Comparison of behavior with SF systems already installed and with lab data obtained in WP2 Origin of data: data analysis report. Expected size of data: some GBs To whom might it be useful? Scientific community, private industrial actors, potential customers
2 FAIR Data	
2.1 Making data findable, including provisions for metadata	Description of the data: Naming will be _WP5_YYYYMMDD_<Producer ID>_xx_field tests
2.2 Making data openly accessible	How will the data be made accessible? Selected relevant data from the in-field operation will be accessible, in relation with those reported in public deliverable D5.1. The user should provide basic contact information, so the MULTIPLHY partners can identify the party of interest. What method or software tools are needed to access the data? Data will be exported as binary data. It will be possible to export part of them in Excel format to be published as open data. How will the identity of the person accessing the data be ascertained? Information about company, activity will be requested during first access.
2.3 Making data interoperable	Public data of the project for this thematic are interoperable by nature
2.4. Increase data re-use (through clarifying licences)	N/A
3 Allocation of resources	All costs related to the data collection and processing are covered by the project budget under WP8

4	Data security	Sensible and non-public data will be stored on a SharePoint available only for the consortium members
5	Ethical aspects	No ethical aspects for MULTIPLHY
6	Other	N/A

5.5. Work package 6: Regulatory framework & guarantee of origin for the hydrogen (CERTIFHY)

DMP component	MULTIPLHY_data_WP6
1 Data summary	Purpose: purchase green electricity, assess green hydrogen certification and issue guarantees of origin via Certifhy platform Data format: .xlsx, .docx Re-use of existing data and if yes, how? Yes, several options available for RE supply; CertifHy methodology as starting point Origin of data: Report on strategy & action plan to supply the project with renewable electricity Report on CertifHy methodology assessment for integrated HTE systems. Expected size of data: MB To whom might it be useful? Potential end-users of THE technology, stakeholders
2 FAIR Data	
2.1 Making data findable, including provisions for metadata	Description of the data: Naming will be _WP6_YYYYMMDD_<Producer ID>_xx_GoO
2.2 Making data openly accessible	Assessment of hydrogen certification while using waste-heat and steam from industrial applications will be made available, in relation to public deliverable D6.2. Regular office software will be sufficient to access to them. How will the identity of the person accessing the data be ascertained? Information about company, activity will be requested during first access.
2.3 Making data interoperable	Public data of the project for this thematic are interoperable by nature.

2.4. Increase data re-use (through clarifying licences)	N/A	
3	Allocation of resources	All costs related to the data collection and processing are covered by the project budget under WP8
4	Data security	Sensible and non-public data will be stored on a SharePoint available only for the consortium members
5	Ethical aspects	No ethical aspects for MULTIPLHY
6	Other	N/A

5.6. Work package 7: Market, techno-economic and environment studies

DMP component		MULTIPLHY_data_WP7
1	Data summary	Purpose: Techno-economic analysis, verification of business cases, investigation of most economical operation modes, elaboration of future roll-out scenarios for hydrogen technologies in refineries, GHG emission avoidance calculations Data format: .xlsx, .docx Re-use of existing data and if yes, how? Yes. Reference data on other systems already developed by SF/PW, and reference data for comparison with other technologies Origin of data: techno-economic calculations Expected size of data: MB To whom might it be useful? Potential end-users of the technology, stakeholders
2	FAIR Data	
	2.1 Making data findable, including provisions for metadata	Description of the data: Naming will be _WP7_YYYYMMDD_<Producer ID>_xx_techno_eco
	2.2 Making data openly accessible	Selected relevant data will be made accessible on the MULTIPLHY website, in relation with public deliverables D7.2 and D7.3. Regular office software will be sufficient to access to them. How will the identity of the person accessing the data be ascertained? Information about company, activity will be requested during first access.
	2.3 Making data interoperable	Public data of the project for this thematic are interoperable by nature.

DMP component		
2.4. Increase data re-use (through clarifying licences)		N/A
3	Allocation of resources	All costs related to the data collection and processing are covered by the project budget under WP8
4	Data security	Sensible and non-public data will be stored on a SharePoint available only for the consortium members
5	Ethical aspects	No ethical aspects for MULTIPLHY
6	Other	N/A

5.7. Work package 8: Dissemination, communication and exploitation

DMP component		**MULTIPLHY_data_WP8**
7	Data summary	Purpose: Insure highest possible impact for MULTIPLHY, Insure the dissemination of the MULTIPLHY results both targeting scientific audience and a broader public audience, managed the produced knowledge generate industrial protection actions and publications. Data format: Formats for web-publicing, pdf, image file formats (e.g. jpg, png etc.), txt, docx, pptx. In general data distributed from WP8 to the public will be web-formats, pdf and similar Re-use of existing data and if yes, how? Yes. Similar exploitation and dissemination data from partners involvement in other research project will be used for inspiration and comparison Origin of data: Processed experimental data from partners, presentation and articles from partners as well as press releases from partners Expected size of data: In the range of MB to GB To whom might it be useful? Scientific community, stake holders from industry and in specific cases also the broader public
8	FAIR Data	
	2.1 Making data findable, including provisions for metadata	Description of the data: Internally in the project the naming will follow the guideline _WP8_YYYYMMDD_<Producer ID>_event For publications, conference presentations etc. journal name and/or conference name will be included in the name along with at least first author name.
	2.2 Making data openly accessible	How will the data be made accessible? Via the projects homepage www.project-multiplhy.cea.fr, via newsletters sent to mailing list, via conference home

		pages (presentations/posters) and scientific journals (open access)
		What method or software tools are needed to access the data?
		No additional software or tools than standard software and tools on laptops
		Have you explored appropriate arrangements with the identified repository?
		Hosting of homepage is arranged and will be hosted by CEA throughout the project
	2.3 Making data interoperable	Public data of the project for this thematic are interoperable by nature
	2.4. Increase data re-use (through clarifying licences)	Raw data gathered in and/or as product of/generated by WP8 will be stored and remain re-usable in form of raw data (draft of publications, press releases, presentations etc.) being stored at the project Sharepoint and stored for 5 years after finalizing the MULTIPLHY project. Re-use of data after end of the project by project members and by third parties will be possible if/when partners provide their written accept for the re-use of data generated as part of WP8, except for data already public
9	Allocation of resources	The costs will be covered by the project budget for this WP. License to sharepoint is covered by CEA. CEA will be responsible for the data management when received from partners and prior to publishing
10	Data security	Data, news, press-release, articles, presentations will be stored on a Sharepoint site only open for project members until agreed to be published CEA ensures safety of data storage for long-term preservation and curation in WP8
11	Ethical aspects	No ethical aspects for MULTIPLHY
12	Other	N/A

Appendix 1: How to fill MULTIPLHY dataset table

Following table summarizes the main issues listed in "H2020 template: Data management plan" that have to be addressed for each work package dataset.

DMP component	MULTIPLHY_data_WPx_YYYYMMDD_<Producer ID>_xx_<Data ID>
1. Data summary	Purpose: What is the purpose of the data and its relation to the objectives of the project? Data formats: Will you re-use any existing data and how? What is the origin of the data? What is the expected size of the data? To whom might it be useful ('data utility')?
2. FAIR Data	Findable, Accessible, Interoperable, Re-usable
2.1 Making data findable, including provisions for metadata	- Is there any protocol or standard identification mechanism that make the data discoverable with metadata, identifiable and locatable Description of the data: - Naming convention, - Versioning management, - Is metadata creation considered? If so, metadata description.
2.2 Making data openly accessible	- Describe the data that will be made openly available. In case of restrictions detail the reasons separating legal and contractual reason from voluntary restrictions. - Specify the repository - Define how data will be accessible (repository, secure platform…) - Define the software or methods needed to access the data and if some specific documentation or procedure is needed. - Can the relevant software be included to the data? - Is an appropriate arrangement concluded with the identified repository - How access will be provided in case of restrictions? - Is there a need of data access committee? - Are conditions for access clear and accessible? - How will be identity of the persons accessing the data be ascertained?
2.3 Making data interoperable	- Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open)

	software applications, and in particular facilitating re-combinations with different datasets from different origins)? - What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable? - Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability? - 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?
2.4. Increase data re-use (through clarifying licenses)	- How will the data be licensed to permit the widest re-use possible? - 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, bearing in mind that research data should be made available as soon as possible. - Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? - If the re-use of some data is restricted, explain why. - How long is it intended that the data remains re-usable? - Are data quality assurance processes described?
3. Allocation of resources	All costs related to the data collection and processing are covered by the project budget with dedicated person months under WP7. - What are the costs for making data FAIR in your project? - How will these be covered? - Who will be responsible for data management in your project? - Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?
4. Data security	- What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? - Is the data safely stored in certified repositories for long term preservation and curation?
5. Ethical aspects	- Are there any ethical or legal issues that can have an impact on data sharing? - Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?
6. Other	N/A

Appendix 2: Detailed analysis for each Work tasks

Task No	Task Title	Lead Participant Short Name	Describe the generated Data	Data provider	File types Examples of type of format: *.xlsx , *.mp3,*.doc,*.pdf, Jpeg, Audio	Openly accessible: Yes No N/A
T2.1	Definition of testing protocol and preparation of test setups	CEA	Harmonized testing protocols for the evaluation of cell and stacks in terms of performance and durability in rSOC	CEA, SF, ENGIE	Docx, pdf	Yes, public deliverable D2.1
T2.2	Short-term release testing and evaluation	CEA	Performance and short term characterization of 3 types of stacks	CEA, SF	.xlsx, .docx, .pdf,.txt, .mat or .csv	No
T2.3	Long-term testing and evaluation for further upscaling	CEA	Durability test for 2 type of stacks	CEA, SF	.xlsx, .docx, .pdf,.txt, .mat or .csv	Yes, final data correspondi ng to data reported in D2.3
T3.1	Preparation for system integration and operation concept	NESTE	Interface specifications are described for process, utilities, signals and electrical. High level operation concept.	Neste, SF, PW	.xlsx, .docx, .pdf	No
T3.2	HTE system engineering and optimization	SF	Elaboration of production strategy, operation strategy,	SF	.xlsx, .docx, .pdf	No
T3.3	Cost analysis and targeting	SF		SF	.xlsx, .docx, .pdf	No
T3.4	Service and maintenance concept	SF		SF	.xlsx, .docx, .pdf	No
T3.5	HTE system manufacturing, commissioning and FAT	SF	Commissioning and FAT Protocol		.xlsx, .docx, .pdf	No, confidential deliverable D3.2
T3.6	Design, manufacturing and commissioning of auxiliary units	Paul Wurth	Design, Commissioning and FAT Protocol	PW, SF, NESTE	.xlsx, .docx, .pdf	No
T4.1	Engineering and preparation of installation site	NESTE	Engineering and execution of site integration scope.	Neste, SF, PW	.xlsx, .docx, .pdf	No
T4.2	Site integration of HTE, HPU and auxiliaries and	NESTE	Site systems/interfaces will be commissioned	Neste, SF, PW	.xlsx, .docx, .pdf	No

	system commissioning, SAT		and tested by Neste, SF and PW prior to the HTE and HPU connection. Followed by SAT.			
T4.3	Validation of operation procedures	NESTE	Off-line testing and connection to on-line operation. Training. The agreed acceptance criteria and protocols will be filled and signed for integration of the units into the refinery.	Neste, SF, PW	.xlsx, .docx, .pdf	No
T5.1	Long-term technology demonstration and assessment	SF	Mainly data of efficiency, availability, performance	NESTE, SF; PW	.xlsx, .docx, .pdf, .txt, .csv	Yes, public deliverable D5.1
T5.2	Measurement and evaluation of the degradation	SF	Steam quality, standard operating points, post-mortem analysis	NESTE, SF	.xlsx, .docx, .pdf, .txt, .csv	No, confidential deliverable D5.2
T6.1	Power purchasing of renewable electricity	engie	Possible options for RE supply	Engie, NESTE	.pdf	No, confidential deliverable D6.1
T6.2	Hydrogen certification via the CertifHy scheme	engie	Assessment of CertifHy methodology for HTE system	Engie, SF	.pdf	Yes, public deliverable D6.2
T7.1	Techno-economic analysis and energy management strategy assessment	SF	TEA Report	SF, NESTE, ENGIE, Paul Wurth	.pdf	Yes, public deliverable D7.2
T7.2	GHG avoidance potential of hydrogen in refining industry	engie	Information related to current process into the refinery, sizing and detailed design including energy consumption and experimental data	NESTE, SF, Paul Wurth	.docx, .xlsx, .pdf, .jpg	Yes, public deliverable D7.3
T7.3	Technical evaluation of hydrogen usage in refining industry	SF	Information related to use cases for refining industry	SF, NESTE, CEA, Engie	.pdf	Yes, public deliverable D7.2
T7.4	Future strategies for intermittent RES	SF				N/A
T8.1	Dissemination, communication and project website	CEA	Communication tools (logo, templates, leaflet, newsletter, website)	all	.docx, .pptx, .pdf, .jpg	yes

			Scientific papers, conference proceedings, posters, flyers			
T8.2	IP management and impact assessment	SF	Report		.pdf	No
T8.3	Exploitation roadmap	SF	Report		.pdf	No