This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 875193

MODALIS 2

MODelling of Advanced LI Storage Systems

DATA MANAGEMENT PLAN (DMP)

This project has received funding
from the European Union's Horizon 2020 research and innovation programme
under **Grant Agreement No. 875193**

Deliverable number: D7.2

Due date: 31.06.2020

Nature[1]: ORDP

Dissemination Level[1]: PU

Work Package: WP7

Lead Beneficiary: IFPEN

Contributing Beneficiaries:

[1] **Nature:** R = Report, P = Prototype, D = Demonstrator, O = Other
Dissemination level PU = Public
PP = Restricted to other programme participants (including the Commission Services)
RE = Restricted to a group specified by the consortium (including the Commission Services)
CO = Confidential, only for members of the consortium (including the Commission Services)
Restraint UE = Classified with the classification level "Restraint UE" according to Commission Decision 2001/844 and amendments
Confidential UE = Classified with the mention of the classification level "Confidential UE" according to Commission Decision 2001/844 and amendments
Secret UE = Classified with the mention of the classification level "Secret UE" according to Commission Decision 2001/844 and amendments

Version	Date	Description
0.1	23.06.2020	First version by IFPEN (M.Petit)
0.2	17.07.2020	Second version completed based on partners contributions
0.3	17.07.2020	Review done by K&S
0.4	24.07.2020	Approval Steering Committee

Executive Summary

Data management guidelines relevant for MODALIS² project are described in order to ensure FAIR Data Management compliant with Horizon 2020 guidelines.

1 Introduction

1.1 Purpose and Scope of the Deliverable

This deliverable aims at providing information on how data, reports and publications are handled within the MODALIS² project.

1.2 Process of Development

This deliverable has been established based on the “Guidelines on FAIR Data Management in Horizon 2020 version 3.0”, following the general rules issued in the Grant Agreement, the Consortium Agreement and the deliverables D7.1 and 7.3.

1.3 References

1.3.1 Internal documents

MODALIS² Grant Agreement, Document Ref. Ares(2019)6410436 - 16/10/2019 (file: Grant Agreement-875193-MODALIS2.pdf on the DMS)

MODALIS² Consortium Agreement - 13/02/2020 (file: [FINAL VERSION] MODALIS 2_Consortium Agreement.pdf on the DMS)

1.3.2 External documents

European Commission Directorate-General for Research & Innovation (26th July 2016) Guidelines on FAIR Data Management in Horizon 2020. Online: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf , consulté le 22 juin 2020.

1.4 Acronyms

2 Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?	The purpose of the data collection is to provide partners with experimental and modelling data in order to design and evaluate the performances of the cells used in the scope of the project, model and validate the models developed.
What types and formats of data will the project generate/collect?	Depending on the source of the data: - *Experimental data*: o text format o csv files o excel files o word o ppt files o Matlab table files (*.mat) o Matlab figures (*.fig) o JSON files - *Models* o Simcenter Amesim ame file o Battery Design Studio, TBM file o STAR-CCM+, sim file o COMSOL multiphysiscs model .mph file o Crystallographic information file (.cif) o text format .cvs and .xls files, and .jpeg or .tif for images o Matlab table files (*.mat), o Matlab figures (*.fig)" o Text format files Quantum Espresso and Crystal outputs
Will you re-use any existing data and how?	Project partners will reuse literature data as well as previously obtained data from partners
What is the origin of the data?	The origin of data in MODALIS² are experiments and modelling from the different tools and software used by partners.
What is the expected size of the data?	The size of data generated depends on the time scale of the experiment and modelling performed (in the range of 1Mo to 10 Go)
To whom might it be useful ('data utility')?	The data can be useful to the battery material, battery manufacturing and battery modelling communities.

3 FAIR data

3.1 Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?	This will be decided along the project.
What naming conventions do you follow?	File names for formal project documents (e.g. Deliverables, Protocols, Periodic Reports) shall appropriately identify purpose, date, content and revision of each document. Therefore, a uniform naming and numbering convention is established within the project as follows: MODALIS²_PROT_200121-20_v03_Meeting1_Lyon.docx Project name; Type of the document; Date; Document version; Document title The above example thus represents: "Protocol/Minutes of the MODALIS² Project in Lyon on January 20-21, 2020 in version 0.3" A detailed explanation of the parts comprising the file name is given below.

Abbrev.	Type of document	Preferred file formats
AG	Agenda for meetings	Word
INFO	General information, messages	Word, Excel, PowerPoint
DRAFT	Drafts, conceptual work	Word, Excel, PowerPoint
LIST	Lists of items	Excel
MAN	Manuals, handbooks, guidelines	Word
MEDIA	Media documents (images, videos, ...)	Various media formats
PR	Presentations	PowerPoint
PROT	Protocol (formal) or minutes (informal) of meetings	Word
REP	Reports (reviews, status/annual reports)	Word

		SPEC	Specifications	Word, Excel
		SCH	Schedules, time plans	Excel
		TEM	Templates	Word, Power-Point, Excel
Will search keywords be provided that optimize possibilities for re-use?	Relevant keyword will be used such as modelling, silicon, NMC… based on careful review of most used keywords on the subject in the literature.			
Do you provide clear version numbers?	Each deliverable version is followed and summarized in a table in its first pages. Each model and dataset will be versioned beginning at number 1 and the version number will be indicated in the end of filename as 'vXX'.			
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.	This will be decided along the project.			

3.2 Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.[2]	Data used in papers and publications will be made openly available. Specific data concerning physical-chemical characterization of material used in the cells will not be shared by default due to intellectual property protection as described in MODALIS² Consortium Agreement.
How will the data be made accessible (e.g. by deposition in a repository)?	Data, literature and publication will be deposited in a centralized and certified portal like ZENODO.
What methods or software tools are needed to access the data?	For experimental data, text files will be used allowing access with any text processing software. Specific models require using the software used to develop them such as Simcenter Amesim of BDS. Information on software needed will be given in a specific file where all file format

[2] Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

	used in the project will be listed along the necessary software to read them.
Is documentation about the software needed to access the data included?	Software documentation is embedded with it.
Is it possible to include the relevant software (e.g. in open source code)?	Where relevant and possible with regards to the property rights developed software will be made available through the open source repository ZENODO.
Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.	Partner K&S operates a public website and a document management system (DMS) for the project. Both can be used for sharing data, metadata and documentation during the project. For data sharing after the project end, the consortium will be exploring other means, in particular with certified open access repositories.
Have you explored appropriate arrangements with the identified repository?	No
If there are restrictions on use, how will access be provided?	No, access will be granted to partners on demand.
Is there a need for a data access committee?	No
Are there well described conditions for access (i.e. a machine readable license)?	No
How will the identity of the person accessing the data be ascertained?	Data can be accessed anonymously (Zenodo)

3.3 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) soft-	Data produced in the project will be produced and used by several partners using different experimental devices and different processing and simulation tools. As a consequence, special care will be taken in order to ensure that data will fit all partners requirements in terms of processing and re-use.

ware applications, and in particular facilitating re-combinations with different datasets from different origins)?	Structural information about materials will be made accessible using the CIF file format (Crystallographic Information File).
What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?	Data format will be defined at the beginning of the experimental campaign.
Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?	Data format will be defined at the beginning of the experimental campaign.
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?	This will be addressed over the course of the project.

3.4 Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?	Open Data generated by the project will be made available under non-contaminating license. Data will be publicly distributed through the project website, CORDIS as well as, when relevant, through the free and commercial software of the MODALIS² partners. The distributed data files will mention the applicable license. Commercial data to be integrated with commercial simulation software will be licensed under the standard Siemens EULA (End User License Agreement, https://www.plm.automation.siemens.com/global/fr/legal/online-terms/eula.html) excluding any exclusivity, so allowing a large distribution of the data to Siemens customers.
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.	Confidential data embargo is 10 years after the end of the project as stated by the consortium agreement.

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.	No
How long is it intended that the data remains re-usable?	No limitation of access is planned for experimental and text based data.
Are data quality assurance processes described?	Data quality assurance is provided through reviewing of MODALIS² deliverable by partners.
Further to the FAIR principles, DMPs should also address:	No

4 Allocation of resources

What are the costs for making data FAIR in your project?	Data sharing will be performed by using MODALIS² website, ZENODO and also by submitting papers in Open Access journals which requires extra costs.
How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).	Those costs are already planned in the WP6: Dissemination.
Who will be responsible for data management in your project?	The project coordinator, IFPEN, is responsible of MODALIS² data management.
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)?	No

5 Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?	All reports and publications generated during the project will be stored on the project DMS, which is secured according to state-of-the-art IT measures. All data is included in weekly backup routines. The project guidelines (D7.1) give clear advice for how to communicate within the project and with external parties, in particular that the DMS shall be used for sharing sensitive information. Data is well structured within the DMS and WP7 is continuously checking and maintaining the data repository.
Is the data safely stored in certified repositories for long term preservation and curation?	All data on the project DMS is included in weekly backup routines. For long term preservation and curation each publication will be also recorded on ZENODO the open database of the EU Open AIRE project.

6 Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).	No
Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?	No questionnaires dealing with personal data are planned over the course of MODALIS² project.

7 Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?	No