



CELL-INTEGRATED SENSING FUNCTIONALITIES FOR SMART BATTERY SYSTEMS
WITH IMPROVED PERFORMANCE AND SAFETY

GA 957273

D7.3. – INITIAL DATA MANAGEMENT PLAN

LC-BAT-13-2020 - Sensing functionalities for smart battery cell chemistries



Deliverable No.	7.3	
Related WP	WP7	
Deliverable Title	Initial Data Management Plan	
Deliverable Date	23-02-2021	
Deliverable Type	REPORT	
Dissemination level	Public (PU)	
Written By	Mattin Lucu (IKE)	12-02-2021
Checked by	Iñigo Gandiaga (IKE)	23-02-2021
Reviewed by	Maaïke van der Kamp (UNR)	17-02-2021
	Prasanna Kadirvelayutham (ABEE)	19-02-2021
Approved by	Iñigo Gandiaga (IKE)	23-02-2021
Status	Final	23-02-2021



Document history

Table 1 indicates the document history, in order to keep track on the different modifications and improvements included during the project.

Table 1. Versioning of the Data Management Plan document

Version	Date	Beneficiary	Author and modification
0	12/02/2021	IKERLAN	Mattin Lucu
0.1	19/02/2021	IKERLAN	Mattin Lucu The document has been updated after the reviewing process by UNR and ABEE.
1	23/02/2021	IKERLAN	Iñigo Gandiaga Last review and submission to EC portal.



Summary

In the recent years, there is a strong incentive from the European Commission (EC) towards the implementation of open access to scientific publications and research data derived from the European Union (EU) funded research projects. This is reflected in the implementation of the Open Research Data (ORD) pilot program (in which SENSIBAT participates), as well as in the efforts to standardise procedures to achieve Findable, Accessible, Interoperable and Reusable (FAIR) research data. The goal behind this incentive is to maximise the impact of EU-funded research projects and accelerate discovery through the interaction of the data produced in several research projects.

This initial Data Management Plan (DMP) aims at proposing a general approach to handle the research data during and after the end of the SENSIBAT project, providing guidelines to specify which data will be collected, processed and/or generated, to what extent this data will be publicly available, and how data will be curated and preserved (including after the end of the project). The document also addresses several aspects on how to make the data produced in the SENSIBAT project as FAIR as possible, following the indications provided by the EC. This deliverable does not include any deviation from the objectives and timings planned in the Grant Agreement of the project.

It is noteworthy that, concurrently, a special effort is made within the Battery2030+ consortium to coordinate the efforts towards a data "FAIRification", among six EU-funded projects. Therefore, and in accordance with the guidelines provided by the EC, this initial DMP is intended to be a living document which will be continuously updated as the implementation of the project progresses and considering the discussions between the SENSIBAT and Batteries2030+ consortia. Accordingly, this document will be periodically revised and completed, and the final DMP will be submitted in M36 of the SENSIBAT project.



Table of Contents

1	Introduction.....	8
1.1	SENSIBAT project summary.....	8
1.2	Towards FAIR data.....	8
2	Data Summary.....	10
2.1	Summary tables for the produced research data.....	10
2.2	Generic metadata.....	11
2.2.1	Data reference.....	12
2.2.2	Digital Object Identifier.....	12
2.2.3	Name.....	12
2.2.4	Purpose of data collection.....	13
2.2.5	Data format.....	13
2.2.6	Expected size of the data.....	13
2.2.7	Repository.....	13
2.2.8	Accessibility level.....	13
2.2.9	Rationale for accessibility restriction.....	14
2.2.10	Embargo period.....	14
2.2.11	Owner.....	14
2.2.12	Lead partner.....	14
2.2.13	Involved partners.....	14
2.2.14	Work-package.....	14
2.2.15	Input to Work-package.....	14
2.2.16	Re-used data.....	14
2.3	Data collection.....	15
3	FAIR data.....	16
3.1	Making the data Findable, including provision of metadata.....	16
3.2	Making the data openly accessible.....	17
3.3	Making the data interoperable.....	18
3.4	Increase data re-use (through clarifying license).....	18
3.5	Allocation of resources.....	19
3.6	Data Security.....	20
4	Discussion and further topics.....	21
5	Conclusion.....	23



6	Risks	24
7	References	25
8	Acknowledgement	26



Abbreviations

Symbol / Abbreviation	
BIG-MAP	<i>Battery Interface Genome-Materials Acceleration Platform</i>
BMS	<i>Battery Management System</i>
CC BY	<i>Creative Commons Attribution license</i>
DMP	<i>Data Management Plan</i>
DOI	<i>Digital Object Identifier</i>
EC	<i>European Commission</i>
EU	<i>European Union</i>
FAIR	<i>Findable, Accessible, Interoperable, Reusable</i>
ORD	<i>Open Research Data</i>
WP	<i>Work-Package</i>



1 Introduction

1.1 SENSIBAT project summary

SENSIBAT's overall objective is to develop a sensing technology for Li-ion batteries that measures in real-time the internal battery cell temperature, pressure, conductivity, and impedance (separately for the anode, cathode, and electrolyte).

The data and insights from these new sensing technologies will be used for the development of improved battery models and state estimator functions based on an improved understanding of how, where and when degradation and failure mechanisms occur. These functions will be included in the Battery Management System (BMS).

SENSIBAT's approach consists of five steps: 1) develop the required battery cell sensor technology, 2) subsequently integrate this sensor technology into 1Ah and 5Ah pouch battery cells, 3) incorporate the 5 Ah cells in a 24V battery module with BMS, 4) use the data from the internal sensing technologies to develop robust and advanced battery models and state estimation functions, 5) carry out a cost-benefit analysis for the batteries with sensors as well as a recycling study of the cells.

Accordingly, different types of data are expected to be produced during the project, coming from the development and integration of battery internal sensors, the experimental cycling tests carried out, the developed battery models and state estimations algorithms, etc.

This initial Data Management Plan (DMP) aims at proposing a general approach to handle the research data during and after the end of the project, specifying which data will be collected, processed and/or generated, to what extent this data will be publicly available, and how data will be curated and preserved (including after the end of the project).

1.2 Towards FAIR data

In the recent years, there is a strong incentive from the European Commission (EC) towards the implementation of open access to scientific publications and research data derived from the European Union (EU) funded research projects.

Accordingly, many recently funded EU research projects take part in the Open Research Data (ORD) pilot program, which implies that, in addition to provide open access to the derived research publications, the provision of open access to the data underlying publications is also promoted. Publishing additional research data is also strongly encouraged, following the approach "as open as possible, as closed as necessary" [2].

The goal behind this incentive is to maximise the impact of EU-funded research projects and accelerate discovery through the interaction of the data produced in several research projects.

Nevertheless, ensuring usefulness of the produced research data beyond the objectives of the single project is a challenging task. As highlighted in [3], «beyond proper collection, annotation, and archival, data stewardship includes the notion of "long-term care" of valuable digital assets, with the goal that they should be discovered and re-used for downstream investigations, either alone, or in combination with newly generated data. The outcomes from good data management and stewardship, therefore, are high quality digital publications that facilitate and simplify this ongoing process of discovery, evaluation, and reuse in downstream studies». In order to achieve this, the concepts of Findable, Accessible, Interoperable and Reusable (FAIR) data raised, reproduced here from [4] for convenience:

«Data are **Findable** when they are described by sufficiently rich metadata and registered or indexed in a searchable resource that is known and accessible to potential users. Additionally, a unique and persistent identifier should be



assigned such that the data can be unequivocally referenced and cited in research communications. The identifier enables persistent linkages to be established between the data, metadata and other related materials in order to assist data discovery and reuse. Related materials may include the code or models necessary to use the data, research literature that provides further insights into the creation and interpretation of the data and other related information.

Accessible data objects can be obtained by humans and machines upon appropriate authorisation and through a well-defined and universally implementable protocol. In other words, anyone with a computer and an Internet connection should be able to access at least the metadata. It is important to emphasise that Accessible in FAIR does not mean Open without constraint. Accessibility means that the human or machine is provided - through metadata - with the precise conditions by which the data are accessible¹ and that the mechanisms and technical protocols for data access are implemented such that the data and/or metadata can be accessed and used at scale, by machines, across the web.

Interoperable data and metadata are described in the FAIR principles as those that use a formal, accessible, shared, and broadly applicable language for knowledge representation. They use vocabularies which themselves follow the FAIR principles, and they include qualified references to other data or metadata. What this describes is semantic interoperability. In other words, the data are described using normative and community recognised specifications, vocabularies and standards that determine the precise meaning of concepts and qualities that the data represent. It is this that allows the data to be 'machine-actionable' so that the values for a set of attributes can be scrutinised across a vast array of data sets in the sound knowledge that the attributes being measured or represented are indeed the same. Interoperability is an essential feature in the value and usability of data. It is not only semantics but also technical and legal interoperability. Technical interoperability means that the data and related information is encoded using a standard that can be read on all applicable systems. In FAIR, legal interoperability falls under the principle that data should be 'Reusable'.

For data to be **Reusable**, the FAIR principles reassert the need for rich metadata and documentation that meet relevant community standards and provide information about provenance. This covers reporting how data was created (e.g., survey protocols, experimental processes, information about sensor calibration and location) and information about data reduction or transformation processes to make data more usable, understandable or 'science-ready'. [...] Open community-endorsed formats also play a key role in reusability. The ability of humans and machines to assess and select data on the basis of criteria relating to provenance information is essential to data reuse, especially at scale. Reusability also requires that the data be released with a 'clear and accessible data usage license': in other words, the conditions under which the data can be used should be transparent to both humans and machines».

That said, one of the main objectives of this initial DMP is also to provide indications on how to make the data generated within the SENSIBAT project as FAIR as possible.

It is noteworthy that, concurrently, a special effort is made within the Battery2030+ consortium to coordinate the efforts towards a data "FAIRification", among six EU-funded projects. The development of vocabulary/ontologies adapted to the battery-related disciplines, the development of other tools for metadata generation are some topics which will be further discussed among SENSIBAT and the other Battery2030+ projects in the following years. Therefore, and in accordance with the guidelines provided by the EC, this initial DMP is intended to be a living document in which information will be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur [1]. Accordingly, this document will be periodically revised and the final DMP will be submitted in M36 of the SENSIBAT project.

¹ «The 'A' in FAIR does not necessarily mean 'Open' or 'Free', but rather, gives the exact conditions under which the data are accessible».

2 Data Summary

This section aims at describing the method which will be used in the SENSIBAT project in order to achieve a comprehensive overview of the data generated during the course of the project.

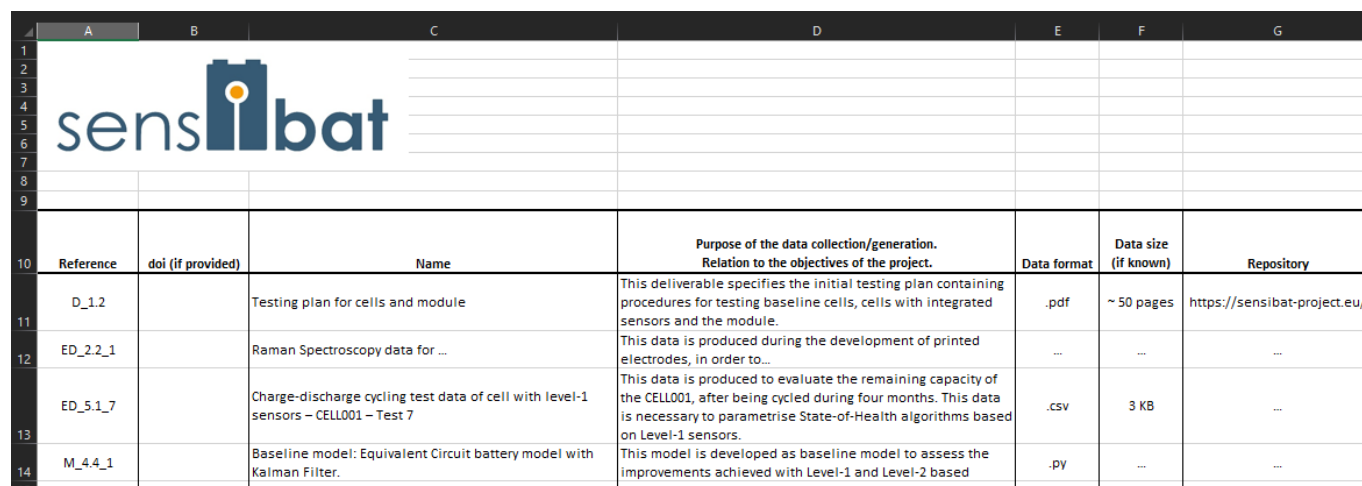
Several aspects are covered, including:

- i) A description and template of the main data summary table, which aims at enumerating all the data produced within the project and providing general information on the purpose, provenance, and accessibility of this data.
- ii) A list and description of the generic metadata required to be completed for all the data.
- iii) Indications on how the produced (meta)data will be collected.

2.1 Summary tables for the produced research data

All the data generated in the project shall be referenced in a spreadsheet called "SENSIBAT_Data_summary" table, which will be completed during the project. This latest version of this spreadsheet can be found on the Mett Platform.

Figure 1, Figure 2 and Figure 3 provide an overview of the template used for this data table.



Reference	doi (if provided)	Name	Purpose of the data collection/generation. Relation to the objectives of the project.	Data format	Data size (if known)	Repository
D_1.2		Testing plan for cells and module	This deliverable specifies the initial testing plan containing procedures for testing baseline cells, cells with integrated sensors and the module.	.pdf	~ 50 pages	https://sensibat-project.eu/
ED_2.2_1		Raman Spectroscopy data for ...	This data is produced during the development of printed electrodes, in order to...
ED_5.1_7		Charge-discharge cycling test data of cell with level-1 sensors - CELL001 - Test 7	This data is produced to evaluate the remaining capacity of the CELL001, after being cycled during four months. This data is necessary to parametrise State-of-Health algorithms based on Level-1 sensors.	.csv	3 KB	...
M_4.4_1		Baseline model: Equivalent Circuit battery model with Kalman Filter.	This model is developed as baseline model to assess the improvements achieved with Level-1 and Level-2 based	.PY

Figure 1. Template for the data summary table, to be completed during the SENSIBAT project, part 1.



	E	F	G	H	I	J	K
1							
2							
3							
4							
5							
6							
7							
8							
9							
10	Data format	Data size (if known)	Repository	Accessibility level	Rationale for accessibility restriction	Embargo period	Owner (individual partner or joint ownership)
11	.pdf	~ 50 pages	https://sensibat-project.eu/	Level 5		none	
12	Level 5		none	
13	.csv	3 KB	...	Level 5		none	
14	.py	Level 5		none	
15							

Figure 2. Template for the data summary table, to be completed during the SENSIBAT project, part 2.

L	M	N	O	P	Q	R
				Re-used data		
Lead partner	Involved partners	In which WP it is produced	To which WP it is communicated	Existing data being re-used	Origin of the data being re-used	Data utility: to whom will it be useful
		WP1	WP5	No		
		WP2		No		
		WP5	WP4	No		
		WP4		No		

Figure 3. Template for the data summary table, to be completed during the SENSIBAT project, part 3.

In addition to enumerate the data produced in the SENSIBAT project, this data table also includes additional descriptive information on the data itself, so-called metadata. Many fields of the metadata shall be intrinsically related to the type of data described (e.g., chemistry of the cell, of the data corresponds to a cell testing data). These fields will be defined case by case in further versions of this DMP.

Nevertheless, a common batch of the metadata shall be defined for all the data, related to data referencing, accessibility, provenance, etc. Section 2.2 provides an initial approach for defining this generic batch of metadata.

2.2 Generic metadata

As highlighted in [1], [3], [4] and in Section 3 of this document, a key and challenging requisite towards the generation of FAIR data is the generation of rich metadata. This metadata should include contextual documentation about the data itself, clearly specifying a persistent identifier for each data, as well as descriptive information about the data.



The definition of suitable metadata is a complex task, and the standards for that are clearly dependent on the research discipline or domain. It is noteworthy that metadata standards are missing for the battery-related research areas and should be further developed by the battery community in the following years (see discussions in Section 4).

In this initial DMP, basic fields are defined for the metadata, which shall be documented for each data generated in the SENSIBAT project using the data summary file on Mett. These basic fields in the data summary file are described below.

2.2.1 Data reference

This first field aims at identifying accurately each data produced within the project. Each data generated within the SENSIBAT project shall be appropriately referenced, and the following reference code will be used:

«*DataType _ WP.Task _ DataNumbering*»

In *DataType*, the type of the data to be referenced will be specified using the following code:

- **D** for deliverables to the European Commission.
- **R** for project internal document, reports.
- **P** for published scientific papers.
- **ED** for experimental data, generated from any kind of laboratory experiments (e.g., electrode printing experiment, cell-level cycling experiment, etc.).
- **M** for developed models and/or algorithms.
- **DEM** for hardware demonstrators.

In *WP.Task*, specify the work-package and task in which the data was generated.

In case more than one data of the same *DataType* were generated in the same task and WP, these will be numbered increasingly, in order to be able to differentiate them. In *DataNumbering*, specify the number of the data generated.

For example, if we refer to a data generated within the fourth battery cycling test carried out in WP5 Task 1, which was e.g., a charge-discharge laboratory experiment, we will refer to it as:

«ED_5.1_4»

If the generated data is e.g., an experimental data (ED) and a deliverable to the European Commission (D) at the same time, always refer to it as an EU deliverable (D).

2.2.2 Digital Object Identifier

Ideally, a Digital Object Identifier (DOI) shall be assigned to all the data generated. For scientific papers, the DOI assigned by the journal will be used. For other data types, the DOI assignment procedure will be further defined during the project (which DOI Registration Agency, procedure, and cost behind this, etc.), in accordance with the Battery2030+ consortium advances in this topic (see discussions in Section 4).

2.2.3 Name

A name shall be assigned to each data generated. The name shall be descriptive about the content of the data, human understandable, and as precise as necessary. For example, if some data corresponds to the variables monitored during a charge-discharge cycle of a cell in which Level-1 sensors were incorporated, a possible name could be:



«Charge-discharge cycling test data of cell with level-1 sensors – CELL001 – Test 1»

In order to progress towards data interoperability and re-usable data, the vocabulary and ontology used in different project should be accorded, and accordingly applied in this field of the metadata. Therefore, further indications on how to complete this field could derive from discussions within the Battery2030+ consortium and further advances in this topic (see discussions in Section 4).

2.2.4 Purpose of data collection

The aim of this field is to specify the reasons behind the data creation within the project, also indicating how this data will contribute to fulfil the objectives of the projects.

2.2.5 Data format

This field specifies the exact format of the data (e.g. .pdf, .csv, .m, .py, etc.)

2.2.6 Expected size of the data

This field specifies, in an approximate way, the expected size of the referred data. Different units could be used to express the size of the data, depending on the data type and format (e.g., number of pages for reports, number of KB for .csv file, etc.).

2.2.7 Repository

The data produced will be stored in a repository, in order to be findable and accessible to others.

The report deliverables to EC are openly accessible on the website of the SENSIBAT project, in the results section (link: https://sensibat-project.eu/sensibat_results/). The Mett Platform will serve as a repository for the EU deliverables reports and other SENSIBAT-documents.

For further data, as well as for deliverables which are not a report, additional repository could be necessary. The repository used to store such data will be defined in the following months between SENSIBAT partners, also considering recommendations from Battery2030+ consortium (see discussions in Section 4).

2.2.8 Accessibility level

According to the EC guidelines towards FAIR data, «The 'A' in FAIR does not necessarily mean 'Open' or 'Free', but rather, gives the exact conditions under which the data are accessible» [4].

In accordance with this, the initial plan includes the definition of several “accessibility levels” for the data generated in the SENSIBAT project, summarised in Table 2.

Table 2. Licence distribution associated with each accessibility levels, depending on project membership.

Accessibility level	License for WP partners	License for SENSIBAT partners	License for Battery2030+ partners	License for everyone
Level 1	All rights reserved.	All rights reserved.	All rights reserved.	All rights reserved.
Level 2	CC BY.	All rights reserved.	All rights reserved.	All rights reserved.
Level 3	CC BY.	CC BY.	All rights reserved.	All rights reserved.
Level 4	CC BY.	CC BY.	CC BY.	All rights reserved.
Level 5	CC BY.	CC BY.	CC BY.	CC BY.

Each accessibility level associates a use license to the data, depending on the membership of the user in different consortia. The acronym “CC BY” corresponds to the “Attribution” license of the [Creative Commons](#) licenses: «This license lets others distribute, remix, adapt, and build upon your work, even commercially, as long as they



credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials. »

As expressed in Table 2, from Level 1 to 5, the accessibility of the referred data increases. Level 1 corresponds to a data which will not be shared at all. Level 2 corresponds to a data which will be shared only with the WP partners. Level 3 corresponds to a data which will be shared only within the SENSIBAT consortium. Level 4 corresponds to a data which will be shared only within the Battery2030+ consortium partners. Level 5 corresponds to public data.

These definitions of the accessibility levels are susceptible to change throughout the duration of the project, as further details shall be defined during the project (refining the exact license types in each level, quantifying the cost associated to formal licensing, if necessary, etc.).

Therefore, the accurate implementation of this table shall be further defined in the following months, in accordance with the “D6.3 Dissemination and Exploitation Plan” of the SENSIBAT project, the Grant Agreement and the Consortium Agreement of the SENSIBAT project, the Battery2030+ consortium recommendations and discussions with the Project Officer.

2.2.9 Rationale for accessibility restriction

In case there is a restriction in the accessibility of the referred data, a rationale for this restriction shall be specified in this field.

2.2.10 Embargo period

In case some embargo period is planned before changing the accessibility level of the referred data, such period shall be specified in this field. Project management team will make sure that the research data is made available as soon as possible.

2.2.11 Owner

This field specifies the owner of the referred data (SENSIBAT partner name). It is possible to define a joint ownership involving several partners.

2.2.12 Lead partner

This field indicates the main partner involved in the generation of the referred data.

2.2.13 Involved partners

If necessary, this field indicates further partners involved in the generation of the referred data.

2.2.14 Work-package

This field indicates in which WP of the SENSIBAT project the referred data is produced.

2.2.15 Input to Work-package

This field indicates to which WP of the SENSIBAT project the referred data is communicated.

2.2.16 Re-used data

In case it is necessary to use additional data from external sources, this shall also be indicated in the data summary spreadsheet document. The following metadata fields shall be then completed.



2.2.16.1 Existing data being re-used

This field indicates by “yes” or “no” if the referred data is re-used from another source external to the SENSIBAT project.

2.2.16.2 Origin of the data being re-used

This field indicates the origin of the data being re-used in the project, preferably by a link to repository for download.

2.2.16.3 Data utility: to whom will it be useful

This field indicates in which WP the referred data will be used.

2.3 Data collection

The data collection method used in the SENSIBAT project will be strongly dependent on the type of the produced data.

The data produced in a report form will be collected using the Mett Platform, as defined in the deliverable “[D7.1 Project Handbook](#)” of the SENSIBAT project. This tool will be used during the project to upload the working materials, temporary drafts and .docx data. Final reports will be available in the project website, in the results section (link: https://sensibat-project.eu/sensibat_results/). The corresponding metadata will be collected manually in the spreadsheet called “SENSIBAT_Data_summary” and introduced in Section 2.1.

The data corresponding to laboratory experimental tests (e.g., cell cycling tests data) will be collected using the data loggers available in each laboratory. Depending on the testing equipment available in the laboratories of the different partners, the format and content of the logged data could differ and will be accorded and specified in the deliverable “D1.2 Testing plan”. Furthermore, such data will be gathered using the Mett Platform. In case they exceed the maximum size permitted by the Mett Platform, additional repositories could be necessary to share the data among the involved partners. This will be further defined among the involved partners, if necessary.

Regarding the topic of metadata collection methods for laboratory tests, this will be further discussed with the partners of the SENSIBAT and Batteries2030+ consortia, in accordance with the guidelines provided in the several workshops planned during the upcoming months.

The data corresponding to models and algorithms (e.g., in code or script formats) will also be collected using the Mett Platform. In case they exceed the maximum size permitted by Mett, additional repository could also be necessary.

Finally, for the developed hardware and demonstrators, their digital existence will be ensured by corresponding descriptive metadata provision and reports, available in the results section of the project website.



3 FAIR data

This section aims at addressing several aspects on how to make the data produced in the SENSIBAT project as FAIR as possible, following the indications provided by the EC in [1].

3.1 Making the data Findable, including provision of metadata

Outline the discoverability of data (metadata provision).

The discoverability of the data is ensured by the provision of metadata. Section 2.1 provides an initial approach to be used in the SENSIBAT project for metadata provision, specifying several fields to be determined. The richness of the metadata is key to achieve FAIR data, and therefore, complementary fields may be created in further versions of the DMP, as new necessities appear. Additional indications could also derive from discussions within the Battery2030+ consortium and further advances in this topic (see discussions in Section 4).

Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

The identifiability of data is handled by fields 2.2.1 “Data reference” and 2.2.2 “Digital Object Identifier” of the metadata.

The referencing code proposed in the Field 2.2.1 “Data reference” aims at ensuring uniqueness of each data identifier within the project. Such referencing code shall also be modified in further versions of the DMP in case inconsistencies appear as the number of referenced data increases.

Digital Object Identifiers shall be assigned at least to each one of the publications derived from the SENSIBAT project. Additionally, as the SENSIBAT project is part of the Open Data Research pilot program, data underlying publication is also promoted, and a DOI shall also be assigned to each one of these data.

The DOI Registration Agency used for that is not defined in this initial DMP. This should be detailed in further version, considering DOI assignment costs (if any) and discussions within the Battery2030+ consortium.

Outline naming conventions used.

The naming convention used particularly influences the field 2.2.3 “Name” of the metadata structure proposed in Section 2.1. This is strongly related to the discipline-dependent ontology and vocabulary developed. An early attempt in the battery field is available in [5], in which some terms are recommended. Additionally, the EU-funded [BIG-MAP](#) project, part of the Battery2030+ consortium, aims at proposing advances on ontology and vocabulary standards for the battery fields. Again, the naming convention used within the SENSIBAT project is planned to be further defined in accordance with the joint propositions from other EU-funded LC-BAT projects in the Battery2030+ consortium.

Outline the approach towards search keyword.

This initial DMP does not include details on any specific approach towards search keywords, as this is strongly related to i) the ontology and vocabulary standards to be further defined, and ii) the data repository and management software used to store and handle the data, which is not still fully defined.



Outline the approach for clear versioning.

In accordance with the guidelines provided by the EC, the Data Management Plan (DMP) is intended to be a living document in which information will be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur [1]. Accordingly, this document will be periodically revised and the final DMP will be submitted in M36 of the SENSIBAT project.

Table 1 indicates the document history, in order to keep track on the different modifications and improvements included during the project.

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.

To this date, there is no widely accepted standards for metadata creation in the battery-field. Section 2.1 indicates an initial approach for metadata provision to be used in the SENSIBAT project. As previously mentioned, more fields are expected to be created during the project as specific needs for additional metadata appear.

3.2 Making the data openly accessible

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so. Specify how access will be provided in case there are any restrictions.

In accordance with the EC guidelines towards FAIR data, «The 'A' in FAIR does not necessarily mean 'Open' or 'Free', but rather, gives the exact conditions under which the data are accessible» [4]. The initial DMP includes the definition of several "accessibility levels" for the data generated in the SENSIBAT project, summarised in Table 2. This information shall be clearly stated for each data, in the corresponding field of the metadata (2.2.8 "Accessibility level"). In case some data would not be openly accessible, the rationale for this shall be clearly stated in the corresponding field of the metadata, see 2.2.9 "Rationale for accessibility restriction".

Additionally, according to the guidelines towards FAIR data, the metadata corresponding to all the data is recommended to be public, even if the data itself is of restricted accessibility. This point will be further discussed during the project within the members of the SENSIBAT and Battery2030+ consortia.

Nevertheless, it is noteworthy that all the EU deliverables of the SENSIBAT project shall be public. All the scientific publications derived from the SENSIBAT project shall be published as open access. Additionally, as the SENSIBAT project is part of the Open Research Data pilot program, research data underlying publications are also promoted to be openly accessible.

Specify how the data will be made available. Specify where the data and associated metadata, documentation and code are deposited.

The report deliverables to EU are planned to be openly accessible on the website of the SENSIBAT project as soon as accepted, in the results section (link: https://sensibat-project.eu/sensibat_results/).

For further data, as well as for deliverables which are not report and could exceed the maximum size permitted by the project website, additional repository could be necessary. The link to this repository will be indicated in the corresponding field of the metadata, i.e., 2.2.7 "Repository".



The repositories used to store such data (and corresponding metadata) will be defined in the following months between SENSIBAT partners, also considering recommendations from Battery2030+ consortium (see discussions in Section 4).

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)?

As the necessity of any specific method or software tool is not clearly defined at this stage of the project, this initial document does not include any detail in this point, which will be addressed in further versions of the DMP.

3.3 Making the data interoperable

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. 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?

The interoperability of the data is strongly related to the discipline-dependent ontology and vocabulary developed. An early attempt in the battery field is available in [5], in which some terms are recommended. Additionally, the EU-funded [BIG-MAP](#) project, part of the Battery2030+ consortium, aims at proposing advances on ontology and vocabulary standards for the battery fields.

Therefore, the vocabulary used within the SENSIBAT project is planned to be further defined in accordance with the joint propositions from other EU-funded LC-BAT projects in the Battery2030+ consortium.

3.4 Increase data re-use (through clarifying license)

Specify how the data will be licenced to permit the widest reuse possible.

Each data will be licensed depending on the assigned “accessibility level”, see section 2.2.8. For the data planned to be openly accessible, a CC BY “Attribution” license of the [Creative Commons](#) licenses shall be used: «*This license lets others distribute, remix, adapt, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.*»

Additionally, in case some data shall be of restricted accessibility, the protection licensing process to be used is not defined yet in this initial DMP. Further details on this shall be decided during the project (assigning the exact license types, quantifying the cost associated to formal licensing, if necessary, etc.). This should be done in accordance with the “D6.3 Dissemination and Exploitation Plan” of the SENSIBAT project, the Grant Agreement and the Consortium Agreement of the SENSIBAT project, the Battery2030+ consortium recommendations and discussions with the Project Officer.



Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed.

The EU deliverables of the project will be available as soon as possible.

The application of embargo periods before data sharing will be specified separately for each data. If an embargo period is necessary, a rationale for this will be clearly defined in the corresponding field of the metadata, see 2.2.9 "Rationale for accessibility restriction". The defined embargo period will be clearly stated in the corresponding field of the metadata, see 2.2.10 "Embargo period".

Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

The usability of the produced data by third parties shall be strongly dependent on the license applied to each data, specified by the defined "accessibility level", see section 2.2.8 and 2.2.9. In case the data is of restricted accessibility, the precise licensing terms shall be defined in accordance with the "D6.3 Dissemination and Exploitation Plan" of the SENSIBAT project, the Grant Agreement and the Consortium Agreement of the SENSIBAT project, the Battery2030+ consortium recommendations and discussions with the Project Officer.

Describe data quality assurance processes.

For the EU deliverables, the quality assurance is based on an internal review process before submission, described in-depth the deliverable "[D7.1 Project Handbook](#)" of the SENSIBAT project.

For further data, the quality assurance is the responsibility of the partners involved in data generation.

Specify the length of time for which the data will remain re-usable.

The length of time for which the data will remain accessible and re-usable shall be defined in further versions of this DMP.

3.5 Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs. Describe costs and potential value of long-term preservation.

The activities related to the data production and curation are part of the person months (and therefore Direct personnel costs) quantified in the Grant Agreement for each partner.

The cost related to long-term preservation (and protection, if necessary) of the generated data is not quantified yet in this initial DMP, and depends e.g., on the data storage repositories used and length of time for which the data shall remain accessible. The costs corresponding to the open access provision of research publications and research data have been included as Other goods and services in the budget.

The first step for total cost quantification is to achieve an exhaustive view of all the data to be produced within the project, which is an ongoing task, and which will be refined throughout the whole duration of the project. The potential value associated with the long-term preservation of data and the costs, including how data will be kept beyond the project, how long and how the costs will be met, will be discussed by the consortium in the coming period. Therefore, more detailed cost estimation will be available in further versions of the DMP.



Clearly identify responsibilities for data management in your project.

The Project Coordinator (IKERLAN) will be responsible for developing and updating the DMP throughout the whole duration of the project, as well as for providing guidelines to the SENSIBAT partners to comply with the DMP. Each partner will be responsible for quality assurance of the generated data and metadata.

3.6 Data Security

Address data recovery as well as secure storage and transfer of sensitive data.

Each consortium partner will take responsibility for proper storage, processing and sharing of the data created during their project activities. Once the research data has achieved its final form, responsibilities for storage and sharing will be defined in accordance with further discussions about the repositories to be used.

As previously stated, the report deliverables to EU are planned to be openly accessible on the website of the SENSIBAT project, in the results section (link: https://sensibat-project.eu/sensibat_results/). These data (i.e., the deliverables) and other project related documents will be stored by Uniresearch using the Mett Platform. Uniresearch has a contract with Mett detailing provisions for data security, continuity, and availability.

For further data, as well as for deliverables which are not a report, responsibilities for data recovery and secure storage will be further defined in accordance with discussions about the repositories to be used.

4 Discussion and further topics

This initial DMP covers several topics on how to handle the data generated from the SENSIBAT project. Nevertheless, many aspects of this management plan are still undefined at this stage of the project. Such aspects will be clarified during the project.

Additionally, in order to gather and leverage the efforts of different EU-funded ongoing LC-BAT projects, many topics deserve further discussion during the different workshops organised among the partners of the Battery2030+ consortium.

In this section, several topics still uncovered, and which need further discussion are listed.

Ontology and vocabulary

To achieve interoperability of the produced data, a common vocabulary needs to be agreed between the different actors of the battery-domain research. This is necessary to correctly define the produced metadata in the project.

The EU-funded [BIG-MAP](#) project, part of the Battery2030+ consortium, aims at proposing advances on ontology and vocabulary standards for the battery fields. This work could serve as a basis for correct definition of the metadata within the SENSIBAT project.

DOI assignment

The research publications derived from the SENSIBAT project, which will be published providing open access, will automatically achieve a DOI from the publisher. Nevertheless, several matters remain unclear regarding the assignment of persistent identifiers to further data. Several questions shall be clarified, among others:

- i) Is it necessary to assign a DOI to all the data made available?
- ii) Which DOI registration agency to use for that?
- iii) Is there any cost to assign and maintain a DOI even after the end of the project?

Repository

The Mett Platform will serve as a repository for the EU deliverables reports. Nevertheless, further type of data (.csv files from experimental tests, etc.) could imply high storage requirements. In such cases, the definition of a distributed repository strategy could be necessary, ensuring at the same time that links to all the data and metadata produced within the project could be accessible in an easy and intuitive way.

Accessibility levels and licensing

In case there is data requiring protection and restricted accessibility for third parties, the precise licensing terms shall be defined in accordance with the "D6.3 Dissemination and Exploitation Plan" of the SENSIBAT project, the Grant Agreement and the Consortium Agreement of the SENSIBAT project, the Battery2030+ consortium recommendations and discussions with the Project Officer.

Licensing terms should be defined considering long-term protection requirements, even after the end of the project.



Improve richness of metadata

Section 2.1 provide basic fields for the metadata provision, common to all the data produced during the project. Nevertheless, depending on the nature of the data, more precise descriptions and information could be required, and additional fields shall be provided.

Additionally, at this stage of the project, the provision of metadata is thought to be done manually, which increases the likelihood of being forgotten or carried out in an inaccurate way. A better way to increase the quality and richness of metadata provision could be to increase its automatization or standardisation, e.g., using for that Graphical User Interface environments to encourage partners to provide metadata in identical way. These topics could be further discussed within the SENSIBAT and Battery2030+ consortia.



5 Conclusion

This document summarises the initial Data Management Plan (DMP) of the SENSIBAT project. A general approach is defined to handle the research data during and after the end of the SENSIBAT project, providing guidelines to specify which data will be collected, processed and/or generated, to what extent this data will be publicly available, and how data will be curated and preserved (including after the end of the project). The document also addresses several aspects on how to make the data produced in the SENSIBAT project as FAIR as possible, following the indications provided by the EC.

It is noteworthy that this initial DMP is intended to be a living document which will be periodically revised and completed, by integrating the ideas raised from further discussions among the SENSIBAT and Batteries2030+ consortia. The final DMP will be submitted in M36 of the SENSIBAT project.

6 Risks

Risk No.	What is the risk	Probability of risk occurrence ²	Effect of risk ³	Solutions to overcome the risk
All WPs.	<i>The work required to provide rich metadata is higher than expected, and the budget necessary exceeds the one initially planned in the proposal.</i>	2 - Medium	2 – Medium <i>The metadata generated is not as rich as necessary and limits the FAIR character of the produced data.</i>	<i>Maximise collaboration with Battery2030+ consortium and particularly BIG-MAP project, which plans high dedication on the development of new tools and ideas in these topics.</i>

² Probability risk will occur: 1 = high, 2 = medium, 3 = Low

³ Effect when risk occurs: 1 = high, 2 = medium, 3 = Low



7 References

- [1] European Commission, "Guidelines on Fair Data Management in Horizon 2020," no. July, 2016.
- [2] European Research Council (ERC), "Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020," no. April, pp. 1–7, 2017.
- [3] M. D. Wilkinson *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship," *Sci. Data*, vol. 3, no. 1, p. 160018, Dec. 2016, doi: 10.1038/sdata.2016.18.
- [4] European Commission, "Turning FAIR into reality: final report and action plan from the European Commission Expert Group on FAIR Data," 2018.
- [5] J. Li *et al.*, "Good practice guide for papers on batteries for the Journal of Power Sources," *J. Power Sources*, vol. 452, no. February, p. 227824, Mar. 2020, doi: 10.1016/j.jpowsour.2020.227824.



8 Acknowledgement

The author(s) would like to thank the partners in the project for their valuable comments on previous drafts and for performing the review.

Project partners

#	PARTICIPANT SHORT NAME	PARTNER ORGANISATION NAME	COUNTRY
1	IKE	IKERLAN S. COOP.	Spain
2	BDM	BEDIMENSIONAL SPA	Italy
3	POL	POLITECNICO DI TORINO	Italy
4	FHG	FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.	Germany
5	FM	FLANDERS MAKE VZW	Belgium
6	TUE	TECHNISCHE UNIVERSITEIT EINDHOVEN	The Netherlands
7	NXP NL	NXP SEMICONDUCTORS NETHERLANDS BV	The Netherlands
8	NXP FR	NXP SEMICONDUCTORS FRANCE SAS	France
9	ABEE	AVESTA BATTERY & ENERGY ENGINEERING	Belgium
10	VAR	VARTA MICRO INNOVATION GMBH	Germany
11	AIT	AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH	Austria
12	UNR	UNIRESEARCH BV	The Netherlands

DISCLAIMER/ ACKNOWLEDGMENT



Copyright ©, all rights reserved. This document or any part thereof may not be made public or disclosed, copied, or otherwise reproduced or used in any form or by any means, without prior permission in writing from the SENSIBAT Consortium. Neither the SENSIBAT Consortium nor any of its members, their officers, employees or agents shall be liable or responsible, in negligence or otherwise, for any loss, damage or expense whatever sustained by any person as a result of the use, in any manner or form, of any knowledge, information or data contained in this document, or due to any inaccuracy, omission or error therein contained.

All Intellectual Property Rights, know-how and information provided by and/or arising from this document, such as designs, documentation, as well as preparatory material in that regard, is and shall remain the exclusive property of the SENSIBAT Consortium and any of its members or its licensors. Nothing contained in this document shall give, or shall be construed as giving, any right, title, ownership, interest, license, or any other right in or to any IP, know-how and information.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957273. The information and views set out in this publication does not necessarily reflect the official opinion of the European Commission. Neither the European Union institutions and bodies nor any person acting on their behalf, may be held responsible for the use which may be made of the information contained therein.