Solid-statE lithium metal bAttery wiTh in situ hyBrid ELecTrolyte

Horizon Europe
THE NEXT EU RESEARCH & INNOVATION INVESTMENT PROGRAMME (2021 - 2027)
HORIZON-CL5-2021-D2-01
DG/Agency: CINEA

HORIZON EUROPE PROGRAMME – HORIZON-CL5-2021-D2-01-03
Advanced high-performance Generation 4a, 4b (solid-state) Li-ion batteries supporting electro mobility and other.

SEATBELT project – Grant Agreement no. 101069726

DELIVERABLE 10.2
Data Management Plan

Funded by
the European Union

History of changes

Version	Publication date	Changes
1.0	31 December 2022	Initial version

Synopsis description

Item	Description	Date (dd.mm.yyyy)
Deliverable No.	D10.2	
Related WP	WP10	
Deliverable Title	Data Management Plan	
Deliverable Date	31 December 2022	
Deliverable Type	REPORT	
Dissemination level	Public (PU)	
Written by	Claire Testori (CNRS)	25.11.2022
Checked by	Cristina Iojoiu (CNRS)	20.12.2022
Approved by	Didier Devaux (CNRS)	21.12.2022
Status	Final of the initial version	

SEATBELT consortium

No.	Participant	Acronym	Type	Country
1	Centre National de la Recherche Scientifique	**CNRS**	RTO	FR
2	Commissariat à l'Energie Atomique et aux Energies Alternatives	**CEA**	RTO	FR
3	Polykey Polymers	**PK**	SME	ES
4	Life Cycle Engineering	**LCE**	SME	IT
5	Centre de Recherches Métallurgiques	**CRM**	RTO	BE
6	Consejo Superior de Investigaciones Científicas	**CSIC**	RTO	ES
7	Blue Solutions	**BS**	IND	FR
8	Münster Electrochemical Energy Technology	**MEET**	UNI	DE
9	Universidad Del Pais Vasco	**UPV**	UNI	ES
10	Zentrum für Sonnenenergie- und Wasserstoff-Forschung Baden-Württemberg	**ZSW**	RTO	DE
11	CIC energiGUNE	**CICe**	RTO	ES
12	Institut Laue-Langevin Europe	**ILL**	RTO	EU
13	Renault	**Renault**	IND	FR
14	Euro Support Advanced Materials	**ES**	IND	NL
15	Imperial College of London	**ICL**	UNI	UK

Table of content

Disclaimer

Funded by the European Union

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them. This project also contributes to the objectives of the Batt4EU Partnership under call topic ID: HORIZON-CL5-2021-D2-01-03 (Advanced high-performance Generation 4a, 4b (solid-state) Li-ion batteries supporting electro mobility and other applications).

1. Overall project presentation

As of 2025, new generations of Li batteries based on silicon/carbon (Gen. 4a) and Li metal (Gen. 4b) anode, where flammable liquid electrolyte is replaced by a non-flammable solid-one, will take over the current Li-ion device. However, only all-solid-state Gen. 4b Li batteries are expected to fulfil the needed cell gravimetric energy density specifications demanded by electromobility and stationary applications. Therefore, SEATBELT ambition is to generate a local EU industry that revolves around a cost-effective, robust all-solid-state Li battery comprising sustainable materials by 2026. SEATBELT intends to achieve the first technological milestone of developing a battery cell (TRL5) meeting the needs of Electric Vehicle (EV) and stationary industry. The low-cost SEATBELT cell is safe-by-design with sustainable and recyclable materials, reaching high energy densities (>380 Wh/kg) and long cyclability (>500 cycles) by 2026 in line with the 2030 EU targets. The cells are produced by low-cost solvent-free extrusion process comprising a combination of innovative materials: thin Li metal, hybrid electrolyte, a safe cathode active material without critical materials and thin Al current collector. The cell design being optimized by interface (operando and atomistic modelling) and process (machine learning) methodologies. In addition, new in situ imaging instrumentation will be developed to investigate safety properties and mechanical deformation to assess cell safety in real conditions. An innovative recycling cycle from materials to cell level will be also established. Thus, SEATBELT will be the start point of a first EU all-solid-state battery value chain, whose main players in Research and Technological Development (RTD) and Industry sectors are within the consortium. So, cells and modules will cycle using industrially relevant protocols dedicated to EV and stationary applications. SEATBELT consortium is composed of 14 beneficiary partners and 3 affiliated entities, and one associated partner, from 7 European countries.

More information at:

Project website: https://seatbelt-project.eu/
CORDIS website: https://cordis.europa.eu/project/id/101069726

2. Deliverable presentation

To ensure a widespread dissemination to the public of the results of the project and, meanwhile, to efficiently protect sensitive / confidential data, a Data Management Plan (DMP) adapted to the SEATBELT project is a necessity. The DMP details the process to produce, gather, collect, and archive the research data by the consortium partners. In SEATBELT, many different types of data will be produced either by experimental apparatus or by data processing and treatment in order to reach the goals of the project related to the battery materials and process development. Therefore, this deliverable is a DMP describing the data management life cycle during and after the project. For this, the DMP template of Horizon Europe project was used and filled using the OPIDoR platform developed by CNRS to specifically address data collection, open-access of data within the FAIR (Findable, Accessible, Interoperable) principle in science. This deliverable presents the initial plan and it will be updated regularly during the course of the project because of the large varieties and amount of data generated.

3. Acknowledgment

The author(s) would like to acknowledge all the consortium partners for their work on this deliverable.

4. Presentation of the DMP

4.1. Overview of the data handling

This deliverable is a plan that established the strategy and the rules to manage and archive all the data collected and generated during the SEATBELT project. The goal is to manage their access to be re-usable in an efficient manner with a defined balance established by the Project General Assembly (PGA) between public and confidential data. This DMP will be updated on a regular basis to take into consideration the increasing amount of data generated and their nature. For the consortium, the goal is to accelerate data knowledge and transmission and fostering of ideas via a specifically designed architecture of the data repository that will be hosted and secured at the Univ. Grenoble Alpes. This repository is an essential part of the DMP and will be used by all consortium partners to upload and download data based on the rules detailed in this deliverable. At last, this repository will be a tool to operate within the FAIR principles when it will come to disseminate data to given target groups, stakeholders, or individuals such as academic scientists. For this, a specific French platform entitled Research Data Gouv. (RDG), created in 2022, will act as an interface between the SEATBELT consortium and the data requester. In short, requested data and open data from SEATBELT will be copied in RDG to enable their access for the data requester while the Univ. Grenoble Alpes repository will be restricted to the consortium partners.

4.2. Presentation of the collected data

In SEATBELT, research data will be provided from practical and modeling experiments. In addition, these data set will be complemented by treatment process, analysis, and comparison between results using series of usual licensed computing software and freeware tools. Thus, the DMP set the ground for the rules to follow the data cycle life from their production to archiving. Data will be managed by collecting the raw data with their corresponding meta-data as well as the processed data to extract meaningful information, their type, size and nature (confidential, public, under embargo prior becoming public), and their naming based on the convention established by the PGA of SEATBELT.

Moreover, the data collection within SEATBELT can be grouped into two categories:

- Personal related data: These data are to be remained confidential such as names, phone numbers, email addresses, affiliation of partners, external individuals, stakeholders, international advisory board, should be restricted to be used by the sole consortium partners.
- Technical related data: All other data falls into this category and are listed into three sub-categories: i) public, ii) sensitive (confidential), iii) under embargo (confidential until release to the public is approved by the PGA). These data can be generated and collected manually or automatically.

4.3. DMP template and structure

For SEATBELT, the DMP is based on the template provided by the European Union, reproduced below, and filled using the OPIDoR platform:

1. Data Summary

- Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

2. FAIR data

2.1. Making data findable, including provisions for metadata

- Will data be identified by a persistent identifier?
- Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
- Will metadata be offered in such a way that it can be harvested and indexed?

2.2.1. Making data accessible : Repository

- Will the data be deposited in a trusted repository?
- Have you explored appropriate arrangements with the identified repository where your data will be deposited?
- Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

2.2.2. Making data accessible : Data

- Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.
- If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- Will the data be accessible through a free and standardized access protocol?
- If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
- How will the identity of the person accessing the data be ascertained?
- Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

2.2.3. Making data accessible : Metadata

- Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?
- How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?
- Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

2.3. Making data interoperable

- What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?
- 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? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?
- Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

Recommandations:

A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)

2.4. Increase data re-use

- How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?
- Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?
- Will the data produced in the project be useable by third parties, in particular after the end of the project?
- Will the provenance of the data be thoroughly documented using the appropriate standards?
- Describe all relevant data quality assurance processes.
- Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

3. Other research outputs

- In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).
- Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

4. Allocation of resources

- What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?
- How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)
- Who will be responsible for data management in your project?
- How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

5. Data security

- What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?
- Will the data be safely stored in trusted repositories for long term preservation and curation?

6. Ethics

- Are there, or could there be, any ethics 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).
- Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

7. Other issues

- Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

4.4. Presentation of the selected repository

All the data generated during the SEATBELT project will be stored and archive at the Univ. Grenoble Alpes academic repository called SUMMER. The data storage space for the project is of 5 TB and will be adapted based on the amount of data collected all over the course of the project. A virtual machine has also been created, entitled WINTER, to secure the access to SUMMER of the consortium partners independently of their location in Europe. This allows them to deposit their files, data and scientific research results related to the project in a secure manner. In addition, this storage space is secured via a specific individual SSH (Secure Shell) key provided to each person working on the SEATBELT project to deposit their data in the most secure way possible. This virtual machine is managed by the coordinator and the project engineer of SEATBELT.

4.5. Presentation of the FAIR data principles

The aim of the FAIR principles is to improve the findability, accessibility, interoperability, and reuse of digital assets. The principles emphasize machine-actionability (*i.e.*, the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention).
The FAIR acronym stands for Findable, Accessible, Interoperable, Reusable, and each of these elements should meet certain criteria described below:

- *Findable*: The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.
- *Accessible*: Once the user finds the required data, she/he needs to know how can they be accessed.
- *Interoperable*: The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.
- *Reusable*: The goal of FAIR is to optimize the reuse of data. To achieve this, metadata and data should be well described so that they can be replicated and/or combined in different settings.

5. SEATBELT initial DMP

Based on the template described provided by the European Commission, the initial version of the DMP of SEATBELT is detailed below. The data are stored in the repository by Work Package (WP) and then decomposed in sub-sections (such as tasks or thematic) and stored using the naming convention established by the PGA. In addition, during the subsequent PGA meeting, update of the DMP will be done to accurately follow the data cycle life (from their production to storage).

5.1. Data summary

- *Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded*

All the data produced for the SEATBELT project are and will be genuinely new. So far, no re-use of data is foreseen for the project except literature data. Indeed, data bibliographic analysis will be thoroughly performed to compare the experimental data produced with state-of-the-art data from literature.

- *What types and formats of data will the project generate or re-use?*

The technical data corresponds to raw file produced by instrumentation machines having their own proprietary type. For example, electrochemical raw files (.mps, .mpt, .zview, etc.) from the manufacturers, physico-chemical characterization (Differential Calorimetry Scanning (DSC), Thermogravimetrics Analysis (TGA), Fourier-Transform Infra-Red (FTIR), Ultra Violet-Visible spectroscopy (UV-Vis), Dynamical Mechanical Analysis (DMA)) are all produced as raw file in their native format. Those files can be only read by proprietary software supplied by the manufacturer of the characterization apparatus. Usually, those files embedded meta-data comprising the machine name and version, controlled and analysis software name and version, the data path onto the computer connected to the apparatus, the file name, the operator's name, the preparation conditions, the technics performed. The meta-data are generated automatically and are partially filled by the operators prior running an experiment. Similarly, modeling data comprises meta-data. Those raw file are then converted into text file readable by any freeware or software with plotting and reading capacities. The meta-data are also embedded in the text file. Moreover, the raw and text files are both saved to ensure data transmission and capability to read all experimental data independently of the proprietary data analysis software. Furthermore, depending on the license of the proprietary data analysis software, the links to download such software will be indicated in the repositories for the consortium partners or the data requester.

The following **Table 1** lists the raw type of data that is and will be generated during the SEATBLET project.

Data type	Example of data extension
Audio	.mp3, .wav, .wma, .ra
Video	.mp4, .avi, .mov, .wmv
Pictures	.jpg, .png, .tif
Documents (e.g. for reports and publications)	.pdf, .pdf/A, .txt, .doc, .docx
Data listing and plots (e.g. spreadsheets and database)	.xls, .xlsx, .op, .csv., .txt
Electrochemical file	.z, .mpt, .mps
Other physico-chemical files	.ta

Table 1. Example of data generated.

In **Figure 1**, the decomposition in a logical way, per Work Package (WP), of the types of data anticipated to be produced during the SEATBELT project is shown. For each WP, the type and the amount of data will evolve along with the maturity of the project.

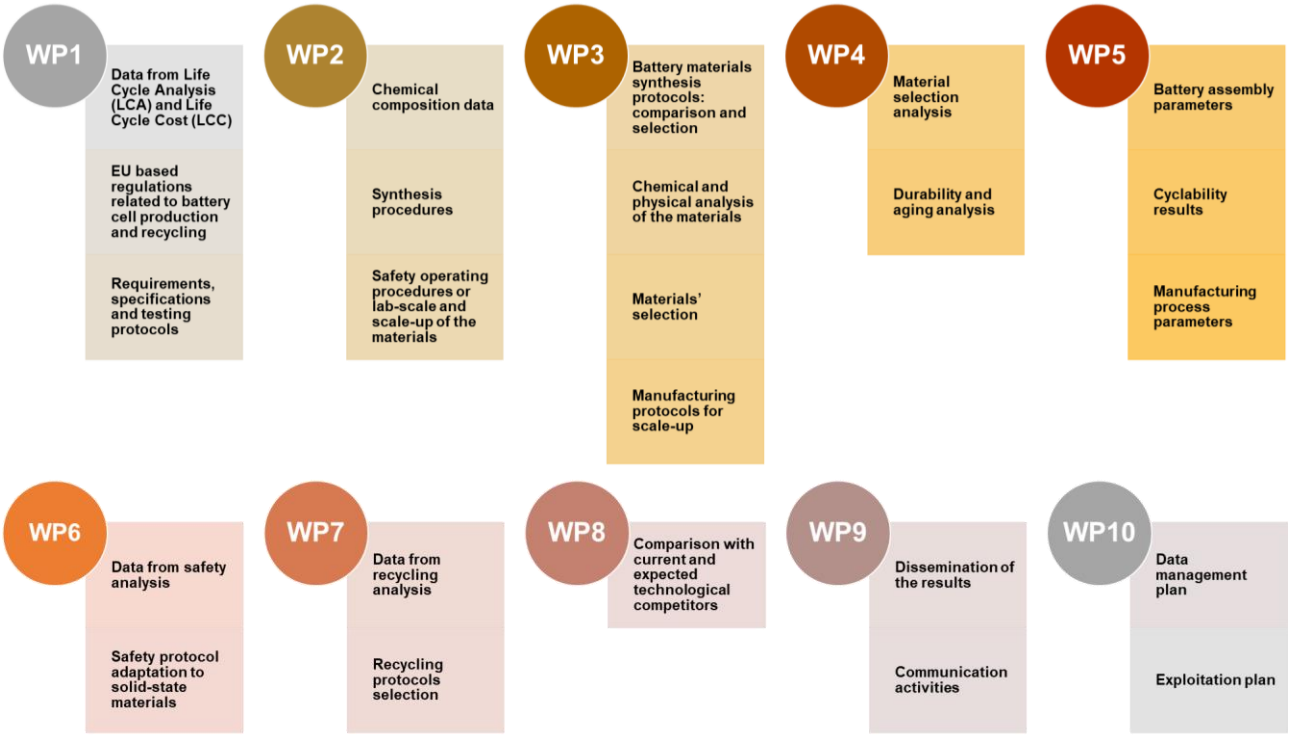

Figure 1. Overview of data produced during SEATBLET project per work package.

- *What is the purpose of the data generation or re-use and its relation to the objectives of the project?*

All the data generated are intended to advance solely the purpose and objectives of the project in order to reach the milestones and develop a reliable solid-state battery cell comprising Li metal as anode and a cathode free of critical raw materials developing a high energy density with long cycle life while being recyclable and sustainable. Therefore, the data generated corresponds to materials' development with their respective manufacturing process, until battery and module assembly conditions and their cycling, safety protocols, electromechanical analysis, as well as the recycling process, and life cycle analysis and life cycle cost. This leads to the production of series of data related to all these aspects so that the project KPIs and milestones are met during the course of the project.

- *What is the expected size of the data that you intend to generate or re-use?*

For physico-chemical and electrochemical data such as battery cycling, the typical size is in the range of hundreds of kilo-octets (Ko) to Mega-octets (Mo). This size range corresponds as well to the size of the file types that are presented in **Table 1**. Moreover, for X-ray and Neutron related data, notably acquired at large-scale facility such as the Institute Laue-Langevin (ILL), one dataset will be in the order of dozens of Giga-octets (Go). Indeed, an imaging data set is based on thousands of projections, each of them being dozen of Mo. Several datasets will be acquired and recorded during the project.

- *What is the origin/provenance of the data, either generated or re-used?*

The data, whose types are listed in **Table 1** and originated form the different works presented in **Figure 1**, are generated by:

- Proprietary analysis apparatus, located at each partners' facility.
- Modeling and code data produced by computing-based technics.
- Computed data by graphical type software.
- Data listing from generated and/or literature data for comparison purpose.
- Reports produced by text type software.
- Audio files for interviews of partner for communication purpose.
- Video files for presentation of the project, partners, and disseminating results.
- Presentation files to gather in documents portions of the obtained results for dissemination purposed (conference, workshops, poster and/or oral presentations), interactions within the consortium partners (restrictive communication intended to remain within the consortium).
- Literature data from database such as scientific one or policies (e.g. EU regulations).

- *To whom might your data be useful ('data utility'), outside your project?*

Outside the project, the data will be useful to anyone involved in the materials and process development of the next-generation of solid-state batteries as well as stakeholders interesting the battery economy (life cycle assessment to gigafactory). Typically, anyone of need to request access to the data produced in SEATBELT will have to read first the DMP so that the data are understandable. Therefore, we expect as primary source of reader:

- The members of the SEATBELT consortium.
- The EU agencies such as CINEA.
- The EU based alliances on battery technology such as BEPA.
- The European commission.
- Stakeholders such as industrials, academics, service providers, policy makers.
- Project consortium and coordinators related to battery technologies.

5.2. FAIR Data

5.2.1. Making data *Findable*, including provisions for metadata

- *Will data be identified by a persistent identifier?*
- *Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*
- *Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?*
- *Will metadata be offered in such a way that it can be harvested and indexed?*

The first step in (re)using data is to find them. Data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services. Therefore, elements from metadata will be used to provide details on the data to help their findability to be later on reuse when needed. Persistent identifiers such as Digital Object Identifier (DOI) will be mainly ascribed to scientific articles published in peer-reviewed journal or in chapter books as well as for large data set such as those produced during large-scale facility experiments, e.g. at synchrotron X-ray facilities where a DOI is automatically associated to the beamtime. Other cases where a persistent identifier may be needed will be assessed during the project by the Project General Assembly (PGA).

For scientific experiments, a raw file is generally created by an apparatus that embedded an entry list. Such list comprises the machine name and version, analysis software name and version, the data path onto the computer, the file name (long and short), the operator's name, the partner acronym, the preparation conditions, and the technics performed. These meta-data are generated

automatically and are partially filled by the operators prior running an experiment. Therefore, data discovery will be based on searching through these extracted meta-data and gathered in a common list file stored in the data storage repository by Work Package number and tasks. In addition, a unique meta-data standard does not exist yet that could combine all the multidisciplinary data types that will be generated during the project. We will rely on the coherence and understandability of the meta-data to navigate by entry-list to make discoverable and findable the data of interest. Therefore, the data will be identified by the various researchers and scientists in the consortium working on the SEATBELT project. They will indicate all the information needed to fully identify data whenever there is a new one.

Keywords will be defined during the project per work package to make data easily findable by the consortium partners. This first version of the DMP does not include specific keywords as those are related to the ontology. In battery science there is no unified description of battery data as detailed in the article of S. Clark et al. in Advanced Energy Materials (May 5, 2022, n° 2102702) detailing that several initiatives are now taking place. Future versions of this DMP will follow the development of these projects to increase the findability of the project data.

5.2.2. Making data *Accessible*

5.2.2.1. Making data accessible: Repository

- *Will the data be deposited in a trusted repository?*
- *Have you explored appropriate arrangements with the identified repository where your data will be deposited?*
- *Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?*

As detailed in section 4.4 (Presentation of the selected repository) on the data will be stored and archive at the SUMMER repository hosted securely by the Univ. Grenoble Alpes. The data storage space for the project is modular, starting with 5 TB and it will be adapted based on the amount of data collected during project lifetime. In addition, a virtual machine was created specifically for the project to secure the access to SUMMER of the consortium partners independently of their location in Europe. This will allow them to deposit their files, data and scientific research results related to the project in a secure manner. In addition, this storage space is secured via a specific SSH key provided to each person working on the SEATBELT project to deposit their data in the most secure way possible. This virtual machine is managed by the SEATBELT coordinator and the project engineer. Moreover, the time of storage of the data takes into account the project time (4 years) as well additional time (at least 4 to 5 years more).

5.2.2.2. Making data accessible: Data

- *Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.*
- *If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*
- *Will the data be accessible through a free and standardized access protocol?*
- *If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?*
- *How will the identity of the person accessing the data be ascertained?*
- *Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?*

The goal is to manage the data access to be re-usable in an efficient manner with a defined balance established by the Project General Assembly (PGA) between public and confidential data. For the consortium, the goal is to accelerate data knowledge and transmission and fostering of ideas via a specifically designed architecture of the data repository that will be hosted and secured at the Univ. Grenoble Alpes. Typically, data reported in official publications will be made available upon request. Any other data will be securely stored in the repository and transmitted to a requester when validated by the PGA and with the possibility for an embargo period to avoid Intellectual Property issue. The embargo period can last couple of months or a year so that a patent application is dully performed.
To access data, and thus disseminate them, a specific French platform entitled Research Data Gouv. (RDG), created in 2022, free for use, will act as an interface between the SEATBELT consortium and the data requester. Requested data and will be first copied in RDG by the project coordinator of project engineer from the Univ. Grenoble Alpes repository then the data requester will be notified along with a protocol to connect and upload the data in RDG. As long as the repository is accessible, data request will be possible, either during and after the project. This procedure to access means a direct contact between the data requester and the project partners which will help to ascertain the validity of the demand which will be ultimately decided by the PGA. The PGA will act as data access committee during and after the project (if possible). The project coordinator will keep contact with the PGA members after the project.

5.2.2.3. Making data accessible: Metadata

- *Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?*
- *How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?*
- *Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?*

The public data will come with the meta-data tagged under Creative Commons CC0. The metadata will contain the information on the procedure to request and access the data.
The storage space for SEATBELT project data is planned for the duration of the project, *i.e.* 4 years plus approximately 5-6 years after the end of the project. The data will thus remain available on this space for approximately 10 years and will be accessible to all persons in the consortium with secure access. Public data will be accessible for an indefinite period of time, *e.g.* the SEATBELT project website. The project coordinator and its institution will ensure data accessibility to external requester. Documentation and references on all the software used to generate and read a data will be provided. When possible, indication on the software license, and the way to acquire the software and its version

used to generate a data will be indicated. In addition, open-source equivalent of proprietary software will be also proposed whenever possible. Moreover, data set will be provided as raw and/or file easily readable, *e.g.* conversion into text file if doable.

5.2.3. Making data *Interoperable*

- *What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?*
- *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? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?*
- *Will your data include qualified references 1 to other data (e.g. other data from your project, or datasets from previous research)?*

In terms of vocabularies, we will first follow the recommendations details in the work of S. Clark *et al.* published recently in Advanced Energy Materials (May 5, 2022, n° 2102702) to ensure that the selected words are understandable by the scientific communities interested in battery materials. This will help to move toward the selection of state-of-the-art practice provided by EU projects to ensure a high level of interoperability.

The technical data are raw file produced by instrumentation machines having their own proprietary type, *i.e.* the native format. Those files can be only read by proprietary software supplied by the manufacturer of the characterization apparatus. Therefore, depending on the license of the proprietary data analysis software, the link to download such software will be indicated to the data requester. Moreover, the raw files will be converted into text file readable by any freeware or software with plotting and reading capacities. The meta-data from the raw files comprising the methodologies, technics, specificities are also embedded in the text files. Consequently, the raw and text files will be both saved onto the repository to ensure data transmission and capability to read all experimental data independently of the proprietary data analysis software and to provide all the information to reproduce the experiments. At last, references to literature works and public data from the project will be found in papers and articles published all along the project.

5.2.4. Increase of data *Reuse*

- *How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?*
- *Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?*
- *Will the data produced in the project be useable by third parties, in particular after the end of the project?*
- *Will the provenance of the data be thoroughly documented using the appropriate standards?*
- *Describe all relevant data quality assurance processes.*
- *Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.*

All necessary documentation will be provided with the raw data to ensure proper interpretation and facilitate data-reuse. As example, machine types, software versions, raw and associated text files will be provided along with experimental conditions, technical details, variable of interest, name of type of the units, etc. In case of uncertainty a project contact such as the project coordinator will take contact with the data requester. Any public data will be re-usable as it will be published in open-access papers and articles with the complete description on the data acquisition and treatment. In line with the obligation set out in the Grant Agreement, the public data from the project, either during

or after the project, will have a CC BY lisence (Creative Commmons) which corresponds to the free use of data as long as the credit and proper reference to the EU funded project is made.
Moreover, for readability a naming convention will be set up for the project that needs to be validated first by the PGA. A special architecture and nomenclature will help to efficiently find the different data deposited on the project storage space.

5.3. Other research outputs

- *In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).*
- *Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.*

Any other research outputs will be taken into consideration during the course of the project and will be fully integrate in future version of the DMP and in the context of the *FAIR* principle.

5.4. Allocation of resources

- *What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?*
- *How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)*
- *Who will be responsible for data management in your project?*
- *How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?*

In terms of cost management, the SUMMER storage space and the WINTER virtual machine from the Université Grenoble Alpes will be covered by the project's eligible credits for a period of at least 10 years. The estimated cost is around 2000 € which is a minor fraction of the total budget. This cost relies on the budget of the coordinator institution, CNRS, which is also responsible for the data management. If needed, the number of years to maintain the repository can be done during and even after the project. The project coordinator will manage the repository after the end of the project. Moreover, ensure publications in open-access journal can be more costly, with an average of 50 papers, at 2000 € per open-access article, leads to a budget that can represent ~ 2 % of the total project budget. Those costs are ensured by each partner.

5.5. Data security

- *What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?*
- *Will the data be safely stored in trusted repositories for long term preservation and curation?*

The selected repository hosted at Univ. Grenoble Alpes is protected by nominative SSH key identification that are provided to each individual in the consortiums, independently of their location in Europe. This key permit to connect to a virtual machine at the interface between the user and the repository. This process enables a high level of security to protect the data generated during the project. The virtual machine is directly managed by the SEATBELT coordinator and the project engineer. Moreover, the data stored onto the repository are duplicate on a daily basis onto a secondary archive managed by Univ. Grenoble Alpes. The data will remain available on this space for at least 10 years (project duration (4 years) + 6 years after the project) and will be accessible to all persons in the consortium with secure access.

5.6. Ethics

- *Are there, or could there be, any ethics 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).*
- *Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?*

SEATBELT activities will be monitored and assessed during the whole project life cycle. For example, ethics checks, reviews and audits will be performed during the project aligned with Horizon Europe Framework Programme Regulation 2021/695: eligible actions and ethical principles and Ethics. The project implementation will not involve research with humans, animal nor the use of personal data. Moreover, the SEATBELT project stands for renewable, clean energies and will strongly contribute to reduction of Greenhouse gases which contributes to global warming and subsequent adverse environmental and human health. The project complies with ethical principles including the highest standards of research integrity as set out in the ALLEA European Code of Conduct for Research Integrity, as well as applicable international and national law, including the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols. Moreover, in the ethic self-statement of the grant application integrity and ethic principles related to data collection are stated. Furthermore, General Data Protection Regulation (GDPR) will be considered to protect personal data and avoid personal data sharing. After data treatment and indexing in the repository, any personal related data will be anonymized. In case, some data indicates personal information the person involved will be aware to accept or deny the release if such information with the right to revoke it at any time.

5.7. Other issues

- *Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?*

So far, no other procedures are foreseen.