

Data management plan



High-Voltage Spinel LNMO Silicon-Graphite Cells and Modules for Automotive and Aeronautic Transport Applications

Horizon Europe | HORIZON-CL5-2021-D2-01-02

*Advanced high-performance Generation 3b (high capacity / high-voltage)
Li-ion batteries supporting electro mobility and other applications*



This project receives funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101069508 (HighSpin).

This publication reflects only the author's view and the European Climate, Infrastructure and Environment Executive Agency (CINEA) is not responsible for any use that may be made of the information it contains.

REVISION HISTORY

Deliverable No.	D8.2	
Deliverable Title	Data management plan	
Deliverable Type	DMP	
Dissemination level	Public	
Written By	Boschidar Ganev (AIT)	28-02-2023
Checked by	Marcus Fehse (CICE), Rasmus Himstedt (CCI), Alexandra Meyer (KIT), Buket Boz (AIT), Cécile Tessier (SAFT)	03-2023
Approved by	Boschidar Ganev (AIT)	14-03-2023
Status	Final	14-03-2023

Version	Date	Who	Change
0.1	28-02-2023	B. Ganev	Populated initial DMP template with responses to all questions
0.2	02-03-2023	B. Ganev	Inclusion of data type updates for WP2 and WP3
1.0	14.03.2023	B. Ganev	Inclusion of further revisions and inputs (mainly to Table 4), finalisation.



TABLE OF CONTENTS

Revision History	2
Project Abstract	5
List of Abbreviations	6
Executive Summary	7
1 Introduction.....	8
2 Data summary	9
3 FAIR data	12
3.1 Making data findable, including provisions for metadata	12
3.2 Making data accessible	12
3.3 Making data interoperable	14
3.4 Increase data re-use	15
4 Other aspects	16
4.1 Other research outputs.....	16
4.2 Allocation of resources	16
4.3 Data security	17
4.4 Ethics	17
4.5 Other issues	17
5 Conclusions	18
6 References	19



LIST OF TABLES

Table 1: List of abbreviations	6
Table 2: Mapping of key DMP sections between this DMP and the reference template.	8
Table 3: Work packages in HighSpin	10
Table 4: Research data expected to be generated or re-used in HighSpin	10
Table 5: Non-research data expected to be generated or re-used in HighSpin	11



PROJECT ABSTRACT

HighSpin aims to develop high-performing, safe and sustainable generation 3b high-voltage spinel LNMO||Si/C material, cells and modules with a short industrialisation pathway and demonstrate their application for automotive and aeronautic transport applications. The project addresses in full the scope of the HORIZON-CL5-2021-D2-01-02 topic, setting its activities in the “highvoltage” line. The project objectives are:

- Further develop the LNMO||Si/C cell chemistry compared to the reference 3beLiEve baseline, extracting its maximum performance.
- Develop and manufacture LNMO||Si/C cells fit for automotive and aeronautic applications.
- Design and demonstrate battery modules for automotive and aeronautic applications.
- Thoroughly assess the LMNO||Si/C HighSpin technology vs. performance, recyclability, cost and TRL.

The HighSpin cell delivers 390 Wh/kg and 925 Wh/l target energy density, 790 W/kg and 1,850 W/l target power density (at 2C), 2,000 deep cycles, and 90 €/kWh target cost (pack-level). The project activities encompass stabilisation of the active materials via microstructure optimisation, the development of high-voltage electrolyte formulations containing LiPF₆ and LIFSI, high-speed laser structuring of the electrodes, and the inclusion of operando sensors in the form of a chip-based Cell Management Unit (CMU).

HighSpin will demonstrate TRL 6 at the battery module level, with a module-to-cell gravimetric energy density ratio of 85-to-90% (depending on the application). Recyclability is demonstrated, targeting 90% recycling efficiency at 99.9% purity. HighSpin aims at approaching the market as a second-step generation 3b LNMO||Si/C in the year 2028 (automotive) and 2030 (aeronautics), delivering above 40 GWh/year and 4 billion/year sales volume in the reference year 2030.



LIST OF ABBREVIATIONS

Table 1: List of abbreviations

Acronym / Short Name	Meaning
DMP	Data management plan
FAIR	Findable, accessible, interoperable, reusable
LNMO	Lithium nickel manganese oxide
Si/C	Silicon graphite
TRL	Technology readiness level
WP	Work package



EXECUTIVE SUMMARY

Data management plans “...typically state what data will be created and how, and outline the plans for sharing and preservation, noting what is appropriate given the nature of the data and any restrictions that may need to be applied”[1].

Good data management can facilitate the sharing of results and knowledge and promote the EU’s open science objectives [2]. Such research data management is often done according to the FAIR principles [3].

The present document provides a first, indicative plan for what kinds of data are expected to be generated in the HighSpin project and how they will be managed, following an established DMP template.



1 INTRODUCTION

The HighSpin project comprises a wide range of activities to achieve its goals (see project abstract, above), including but not limited to key activities such as materials synthesis and characterisation, slurry recipe development, electrode coating, cell assembly, module design and assembly, cell integration, cell and module testing, lifecycle analysis, and recycling investigations. Various types of data will be generated in the course of these activities.

The present document is the first version of the HighSpin data management plan (DMP). It is an indicative plan as to what kind of data the project beneficiaries expect to generate in the course of the project, and how these data will be managed. In establishing the DMP, we have followed a standard DMP template [4].

Section 2 DATA SUMMARY and section 3 FAIR DATA follow the structure of the DMP template. Sections 4-6 of the template are summarised as sub-headings in section 4 OTHER ASPECTS, as shown in Table 2. Green text represents the verbatim questions/prompts from the DMP template whereas the black text underneath is the direct response to these questions.

Table 2: Mapping of key DMP sections between this DMP and the reference template.

This document	DMP template
1. Introduction	-
2. Data summary	1. Data summary
3. FAIR data	2. FAIR data
4.1 Other research outputs	3. Other research outputs
4.2 Allocation of resources	4. Allocation of resources
4.3 Data security	5. Data security
4.4 Ethics	6. Ethics
4.5 Other issues	7. Other issues
-	

The DMP may be revised or updated at any time the consortium deems appropriate during the project, and in any case a final, updated version must be submitted by the end of the project (project month M48), as D8.2 DATA MANAGEMENT PLAN (UPDATED).



2 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.

HighSpin will re-use as needed some data from its predecessor project 3beLiEVe [5]. The reason for reuse is that HighSpin is intentionally conceived as a successor project and expressly aims to re-use and build on previously developed elements (where applicable and useful) to avoid spending time and resources developing them again.

Some of the main types of data that will be re-used include:

- automotive requirements
- data about battery active and passive materials, slurry formulations and processing parameters
- Cell geometry for the automotive cells.

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

A wide range of data types and formats are expected to be generated in the project. While these cannot and need not be taxatively enumerated ex ante, some general remarks can be made. The data types will vary according to the source application in which they are generated. Generally, the consortium will endeavour to utilise common data formats. In some cases, particular laboratory equipment or software may output data in proprietary formats, but where sharing of these data with other partners are necessary a conversion to or a summary in common formats may be possible. Table 4 gives a summary of the expected types of data expected to be generated in the project (Table 3 summarises the project work packages)

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

The purpose of the data generation is to support the activities of the project in the pursuit of project goals. Data about materials, recipes, processes, demonstrator designs, testing, assessment and evaluation results is an integral part of the project activities and needed to determine whether the project has succeeded in its aims, as well as to support dissemination activities.

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

To be determined.

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

Most of the data in the project will be generated in-project. For re-use, see further above.

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



The consortium expects some of the data to be of interest to researchers in the domains of battery and materials research, including modelling, algorithm training, performance benchmarking, manufacturing process research, and others.

Table 3: Work packages in HighSpin

WP	Name
WP1	Identification of the end-user requirements
WP2	Materials development
WP3	Electrode and cell upscaling and engineering (pilot line)
WP4	Automotive: cell and module manufacturing
WP5	Aeronautic: module design, cell and module manufacturing
WP6	Testing and performance qualification, TRL/MRL ex-post assessment, recycling demonstration, LCA/LCC, 2nd life
WP7	Dissemination, communication, exploitation and IPR management
WP8	Project management

Table 4: Research data expected to be generated or re-used in HighSpin

	Type of data	Data formats
WP1	Requirements for automotive and aeronautic applications	Lists and tables (e.g., .xlsx), reports (.pdf)
WP2	Material characterization data; slurry recipes and preparation parameters; processing parameters for electrode manufacturing; electrochemical cycling data	Images (e.g. .jpg, .png, .pdf etc.), data tables (e.g., .txt, .csv, .xy, .xlsx, .opju), lists, recipes and instructions, reports (e.g. pdf, .csv, .docx), video files (e.g., mp4)
WP3	slurry recipes and preparation parameters; processing parameters for electrode manufacturing (including laser patterning); data from post-mortem analysis; electrode designs	Images (e.g. .jpg, .png, .pdf etc.), data tables (e.g., .txt, .csv, .xy, .xlsx, .opju), reports (e.g. .pdf, .csv, .docx), video files (e.g., .mp4)

WP4	Specifications and drawings for automotive cell design	Images (e.g. .jpg, .png, .pdf etc.)
WP5	Data from cell and module design – specifications, drawings	Images (e.g. .jpg, .png, .pdf etc.), CAD files (e.g., .stl)
WP6	Data from cell and module testing; LCA data (e.g., energy and materials use)	Images (e.g. .jpg, .png, .pdf etc.), data tables (e.g., .txt, .csv, .xy, .xlsx, .opju), reports (e.g. .pdf, .csv, .docx), video files (e.g., .mp4)

Table 5: Non-research data expected to be generated or re-used in HighSpin

	Type of data	Data formats
WP7	Website traffic; audience engagement statistics from social media; photos and/or videos of project participants	Exports from (Google) website analytics tool – probably .xlsx Exports from LinkedIn analytics – probably .xlsx and/or .pdf
WP8	Project management data (e.g. financial, resources)	EMDESK exports - Excel lists (.xlsx) and reports (.pdf)



3 FAIR DATA

3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Will data be identified by a persistent identifier?

Unless any project beneficiary applies persistent identifiers for their own use, no persistent data identifiers are envisaged for the purposes of information sharing between partners in the project. However, any datasets made available as open data will leverage the possibilities for persistent identifiers as provided by the host platform (e.g., Zenodo).

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.

The consortium is aware of ongoing efforts to define a common battery language¹, however, not of any ready-to-use taxonomy or similar specifically tailored for batteries. Again, HighSpin consortium will leverage whatever possibilities are offered by the repository in which any open data from the project is hosted. Absent a framework for populating metadata, the beneficiaries providing the data will provide metadata such that interested parties will gain an understanding of the provenance and meaning of the data that is sufficient to interpret and reuse it. Some of the main types of metadata that are envisaged include:

- Description of the test setup from which the data was generated
- Meaning of the different fields/variable in the data files
- Reference to the project in which the data was generated, for further information and context.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

To be determined.

Will metadata be offered in such a way that it can be harvested and indexed?

If the host repository supports this, then yes (provided this can be done with reasonable effort on the part of the consortium).

3.2 MAKING DATA ACCESSIBLE

Repository: Most probably zenodo.org

Will the data be deposited in a trusted repository?

¹ E.g. BattInfo battery interface ontology as part of BIG-MAP project: <https://www.big-map.eu/dissemination/battinfo>



Yes, see above (if data is published in open access).

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

No, we expect that (a) user account(s) will suffice to publish the data.

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

To be determined.

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.

Data are subject to a case-per-case evaluation to determine whether they should be made openly available, in order to protect industrially/commercially sensitive information. Only selected data will be shared after internal consultation and approval in the consortium.

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.

No ex-ante policy on this; to be determined on a case-by-case basis if needed, in conjunction also with activities under *T7.3 Exploitation*, *T7.4 IPR management* and *T8.3 Quality, data and risk management*.

Will the data be accessible through a free and standardized access protocol?

Open data will be published in common formats where this is possible and makes sense. The access protocol will most likely simply be https (we expect to use zenodo.org)

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

To be determined on a case-by-case basis.

How will the identity of the person accessing the data be ascertained?

For open data, no tracking of people accessing the data is planned from HighSpin consortium side. From host platform side, identification will take place presumably via the authentication protocols and server logs of the platform itself.

Is there a need for a data access committee (e.g., to evaluate/approve access requests to personal/sensitive data)?

No personal or sensitive research data is planned to be generated, much less shared, in the HighSpin project. For non-research personal data, beneficiaries in the project are



requested to fill out a consent form regarding the use of photo or video material featuring still or animated images of individuals – in conformity with data protection regulation.

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?

To be determined on a case-by-case basis. Generally, where data is made available, it will contain the information needed in order to access it (we expect this will be simply via zenodo.org).

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

This depends on the policies of the host platform. Once approved for publication as open data, the consortium does not envisage an expiry date. Rather, potential users can determine whether the data is of interest based on its content and age.

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)?

Code generation is not required or expected as part of the project. We expect any data that is generated to be readable with generally available tools (word processors, PDF and image viewers, spreadsheet software).

3.3 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?

Please see section 3.1 and footnote about BattInfo.

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?

No generation of ontologies is planned in this project. We will monitor if/as needed the evolution of BattInfo (<https://www.big-map.eu/dissemination/battinfo>). Any open data shared will be accompanied by adequate metadata to enable interpretation and reuse of the data.

Will your data include qualified references² to other data (e.g., other data from your project, or datasets from previous research)?

² 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



To be determined.

3.4 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.)?

For any open data, this will be provided along with the data (probably as part of the metadata) on the chosen repository.

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?

This will be decided on a case-by-case basis. Generally, where the consortium does decide to make data available in open access, this will be in the public domain with the most permissive possible reuse licences (while respecting project beneficiary IPRs).

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Yes for any data published as open data, since the purpose of such publication is the use and reuse by other parties. For any other data, that is regulated in the project's Consortium Agreement ("Results").

Will the provenance of the data be thoroughly documented using the appropriate standards?

Yes, where standards exist. Where useable standards are absent, adequate documentation with reasonable effort will be produced such that third parties can interpret and reuse the data.

Describe all relevant data quality assurance processes.

Data quality assurance generally rests with the project beneficiary generating the data. For open data, this will be reviewed prior to a publication by a project body to be determined (e.g., WP leaders, individual partners, or as otherwise appropriate). For recording of cell testing results (e.g., cycling), a common framework for use within the project may be adopted or developed.

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.

More information on research outputs can be found in HighSpin deliverables *D7.1 Dissemination and communication plan* and *D8.1 Project handbook*. Both are public deliverables and should be automatically available at <https://cordis.europa.eu/project/id/101069508> once they are approved and released by the project's funding agency.

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/>)



4 OTHER ASPECTS

4.1 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.).

The development of new materials and testing of combinations of such materials in battery cells (and larger aggregations, such as in modules) is a central goal of the project. The management of such outputs is covered through the intellectual property management, regulated in the project's Consortium Agreement and supported and further managed as part of the project's exploitation and IPR management (*T7.3 Exploitation* and *T7.4 IPR management*), and documented in the project's technical deliverables.

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.

A very large proportion of the project's deliverables are classified as "public" with respect to their dissemination level. They will be available at <https://cordis.europa.eu/project/id/101069508> once they are approved and released by the project's funding agency, as well as on the project's website.

4.2 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.)?

To be determined. Generally, the generation of data will be a part of the project activities, therefore no specific/additional resourcing is allocated for this. Where curation and preparation of open data is concerned, this will be done using the resources for the beneficiary generating that data, with support from WP leader and/or project coordinator as needed.

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)

See above. No additional costs for making data FAIR are foreseen at this point.

Who will be responsible for data management in your project?

Each beneficiary manages their own data. Where data is shared to facilitate the project activities, it is the responsibility of each WP leader to keep a good order ("housekeeping") with regard to the data needed for the activities of the WP for which they are responsible. For open data, once it is published to the repository of choice, we will depend on the data management procedures of the repository.



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)?

Internally with beneficiaries – as per their own policies. For open data, according to the policy of the repository.

4.3 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)?

Data security is as per the policies and procedures of each project beneficiary, and particularly those generating or owning the data. Generally, state-of-the-art data security is assumed to be in place. For open data, the consortium relies on the data security provision of the host repository.

Will the data be safely stored in trusted repositories for long term preservation and curation?

Yes, open data, if any, will be stored in a trusted repository (probably zenodo.org).

4.4 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).

No; ethics screening was done at the project proposal stage and in the grant agreement preparation of the project. No noteworthy ethics issues (besides standard health and safety considerations relating to materials handling and exclusion of potential dual-use of the materials) are identified.

Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?

Informed consent for personal data where it regards photos or videos is handled through the consent form as part of WP7. For preservation of these data, where project beneficiaries are the data controllers, these are handled in conformity with the data protection statement (as referenced in the consent form).

4.5 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)?

Not beyond what has been described above on a project level, unless specific beneficiaries wish to do so.



5 CONCLUSIONS

This DMP has disclosed the plans for data management according to best available information at project month M6, following a structured DMP template.

While beneficiaries are encouraged to look for ways to share data with the research community, they are under no obligation to disclose data if this goes against their interests.

An update of this DMP will be published latest at project month M48 (if not before) as D8.2 DATA MANAGEMENT PLAN (UPDATED).



6 REFERENCES

- [1] Dr. Tomasz Miksa, 'Data Management Plans'. SBA Research & TU Wien. [Online]. Available: https://www.ffg.at/sites/default/files/dmp_ffg_miksa.pdf
- [2] 'Open Science', Feb. 10, 2023. https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en (accessed Feb. 17, 2023).
- [3] 'FAIR Principles', *GO FAIR*. <https://www.go-fair.org/fair-principles/> (accessed Mar. 07, 2023).
- [4] 'Horizon Europe - Data Management Plan Template'. May 05, 2021. Accessed: Mar. 07, 2023. [Online]. Available: <https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf>
- [5] 'Delivering the 3b generation of LNMO cells for the xEV market of 2025 and beyond | 3beLiEVe Project | Fact Sheet | H2020 | CORDIS | European Commission'. <https://cordis.europa.eu/project/id/875033> (accessed Mar. 07, 2023).