



Initial Data Management Plan

Deliverable Number D7.2
Deliverable Type R – Document, Report
Dissemination Level PU (Public)
Deliverable Responsible Alexander Thaler, VIF
Document Version & Status V1.0 | Final

Project Acronym greenSPEED
Project Title Green and Sustainable Processes
for Electrode Production
Grant Agreement Number 101069528
Project Coordinator Virtual Vehicle Research GmbH
Project Website www.greenspeed-project.eu



Author(s)

| Name | Organisation | Name | Organisation |
|------------------|--------------|---------------|--------------|
| Alexander Thaler | VIF | Medina Ćustić | VIF |

Reviewers

| Name | Organisation | Date |
|--------------------|--------------|------------|
| Kerstin Reinprecht | VIF | 2022-12-16 |
| Anton Kaifel | ZSW | 2022-12-19 |

Change History

| Version | Date | Name/Organisation | Description |
|---------|------------|--|--|
| V0.1 | 2022-11-25 | Medina Ćustić, VIF | First Draft |
| V0.2 | 2022-11-30 | Alex Thaler, VIF | Structure & Content added |
| V0.3 | 2022-12-01 | Alex Thaler, VIF Medina Ćustić, VIF | Update Chapter 6, 7 and 8 |
| V0.4 | 2022-12-02 | Alex Thaler, VIF Medina Ćustić, VIF | Update Conclusion and References |
| V0.5 | 2022-12-05 | Alex Thaler, VIF Medina Ćustić, VIF | Finalisation of document, conclusion added, update list of abbreviations |
| V0.6 | 2022-12-20 | Alex Thaler, VIF Medina Ćustić, VIF | Update based on reviewer comments |
| V1.0 | 2022-12-21 | Alex Thaler, VIF Medina Ćustić, VIF | Final Version available |

Table of Contents

- 1. Executive Summary 5
- 2. Introduction 6
- 3. greenSPEED Data Management Workshop 7
- 4. Data Summary 8
- 5. FAIR Data 9
 - 5.1 Making data findable, including provisions for metadata..... 9
 - 5.2 Metadata Pre-definition 9
 - 5.3 Making data accessible10
 - 5.4 Making data interoperable11
 - 5.5 Increase data re-use11
- 6. Allocation of Resources.....12
 - 6.1 Costs for making data FAIR12
 - 6.2 Responsibilities of data management.....12
 - 6.3 Long-term preservations12
- 7. Data Security, Ethics and Other Issues13
 - 7.1 Data Security.....13
 - 7.2 Ethics13
 - 7.3 Other Issues.....13
- 8. Conclusion14
- 9. Abbreviations15
- 10. References.....16

List of Figures

Figure 3-1: Agenda of the Data Management Workshop 7

List of Tables

Table 5-1: Naming Conventions 9

1. Executive Summary

greenSPEED is a European funded project under the Horizon Europe programme, HORIZON-CL5-2021-D2-01-04.

The greenSPEED project offers solutions for new sustainable electrode and cell manufacturing processes with reduced energy consumption, lower carbon footprint and ZERO Volatile Organic Compounds (VOCs) emissions. To that aim, the project main target is developing a battery cell comprised of electrodes manufactured by innovative dry processes. Our composite cathode, based on Ni-rich NMC, is to be manufactured by scalable roll-to-roll dry electrode coating process, that fully removes the use of casting-solvents and eliminates the need of energy-intense drying-, condensate and transportation process required in state-of-the-art electrode fabrication.

The greenSPEED high-capacity pure-silicon anode is to be manufactured taking full advantage of our innovative process based on Microwave-Assisted Plasma Enhanced Chemical Vapor Deposition (MW-PECVD), which deposits porous silicon directly on the copper current-collector starting from locally produced silane gas (SiH₄). Moreover, the use of advanced modelling and simulation techniques including digital twins, artificial intelligence, and machine learning are to be employed to predict and optimise cell performance in early development stages, support the cell production process by virtually assessing the influence and importance of production parameters and thus minimising the number of experiments and to accelerate electrode production optimisation steps. The greenSPEED cell aims at increasing energy density (+69%) while reducing energy consumption (-32%) and costs (-21%) of production as compared to state-of-the-art Li-ion cells. The concepts here proposed have been already demonstrated at TRL 2/3 with the aim of reaching TRL 5/6 by the end of the project.

This document is deliverable “D7.2 Initial Data Management Plan” of the greenSPEED project. It is the first version of the Data Management Plan (DMP) of greenSPEED that is planned for project month 6 (December 2022). The DMP is based on the Horizon Europe Programme Guide [1] as well as the Data Management Plan Template [8] and includes all datasets identified at the beginning of the project. It describes the types of data that will be generated and/or gathered during the project, the standards that will be used, the ways in which data will be exploited and shared (for verification or re-use), and in which way data will be preserved. The data management plan is a living document that will be updated on a regular basis.

Within greenSPEED, the coordinator as a person also takes the responsibility of a data manager (DM). The DM defines and monitors the data management mechanisms that are linked to all data requirements, collection, recording and organisation, storage, retrieval, consultation and use, adaptation and alteration, permissions, and access.

The data management within greenSPEED follows the relevant EU and national legislation on data protection. All project partners shall ensure that the EU (e.g. the GDPR (Regulation (EU) 2016/679) and national regulations are complied with.

2. Introduction

greenSPEED is a technology driven project, where data will be collected, processed, and generated. Data is not in the centre of interest of the project but supports the overall process and synchronises the information between the work packages and partners. The proposed actions for the overall data management process within greenSPEED will be performed and monitored in the course of WP7, Task 7.3 Research Knowledge and Data Management Mechanism (led by VIF). It relates to all greenSPEED objectives. The partners are well experienced in FAIR¹ data principles and will further elaborate on these aspects in the respective deliverables and reports.

Because of the competitive industrial nature of greenSPEED, not all data generated by the project can be made available to the public without compromising the legitimate interests of the partners. Nevertheless, data management within the project has a very high priority and will therefore still be carried out. The handling of sensitive data in the project is further regulated in the Consortium Agreement [3] and Grant Agreement [4] of greenSPEED. The data management process is overseen by the project coordinator (VIF).

Within the greenSPEED project, no processing of personal data is planned, except for those necessary for the administration of the project, e.g. contact details. Should this change in the course of the project, appropriate measures will be taken in due time. As the Data Management Plan (DMP) is a “living document”, it will be updated on a regular basis. In any case, the project partners will observe and comply with all applicable legal provisions and regulations regarding data protection, in particular the GDPR (Regulation (EU) 2016/679), also within the framework of data management.

In a first step, the data management process for greenSPEED was defined during the project preparation phase. For D7.2, the data management plan will be further elaborated. This data management plan follows the structure of the Horizon Europe Data Management Plan Template [8] and further defines the methods and standards to be applied, the handling of data, the types of data, open access and the FAIR principle. In addition, data security and protection of data is included.

This DMP has been further created based on the data management workshop conducted with the project partners to ensure that the data management process put in place is appropriate for the project and that the data is processed as efficiently as possible to ensure the quality and sustainability of the results.

¹ FAIR data is data that is **findable, accessible, interoperable**, and **re-usable** [1]

3. greenSPEED Data Management Workshop

A data management workshop has been held to set up the initial DMP with the consortium to ensure the efficient data management within the greenSPEED project as well as the sustainability and quality of the results. For this, the partners joined the online workshop to discuss various approaches on data management. The agenda for the workshop is shown in Figure 3-1.

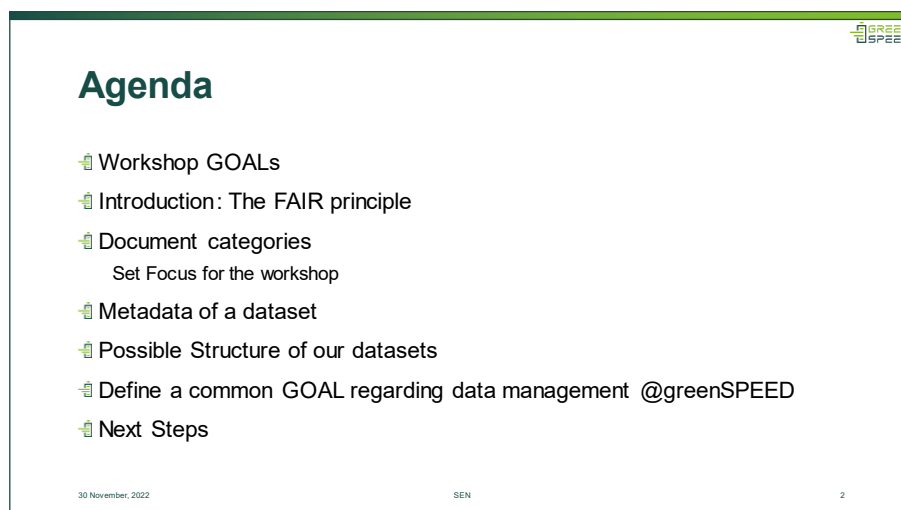


Figure 3-1: Agenda of the Data Management Workshop

During the workshop, the partners have been split in groups to discuss datasets for the project. The following sub-team categories have been chosen:

- Anode datasets
- Cathode datasets
- Cell datasets

The participants have been asked to fill in an Excel file with the following columns:

- Type of dataset
- Responsible partner
- greenSPEED cell generation: Gen0/Gen1/Gen2/all Gen
- Cell integration level: anode/cathode/cell
- Data relation: function/process/material
- Dataset name
- Data format description
- (Expected) amount
- Management system
- Confidentiality level of datasets
- Further additional classification

Please refer to Chapter 5 for further information on how data classifiers are used to elaborate on the FAIR principle.

4. Data Summary

Following greenSPEED's vision of achieving European leadership in battery production with lower carbon footprint, data generated within the project play an important role. In general, data is categorised in one of the two data categories:

- Primary Data: New data generated within the greenSPEED project (e.g. generated from simulation activities, communication activities etc.).
- Secondary Data: Already existing data that are relevant for the greenSPEED activities (e.g. data from open access data bases, data from former projects etc.).

This is especially relevant for the datasets generated within greenSPEED (see Chapter 5).

Deliverables

Project deliverables will be made available as PDF file (approximately 500 MB) and will be based on the Description of Action outlined in the GA [4]. After being accepted by the funding authorities, public deliverables will be available for download via the greenSPEED website ("Results & Media"). Deliverables marked as "sensitive" will not be shared with the public. If agreed by the partners, the executive summary will be made available on the website. All deliverables are stored in the greenSPEED SharePoint where the partners are able to access the documents.

Promotional Material and Videos

Public material to promote the project activities and raise awareness such as flyers, public presentations, and other materials to disseminate greenSPEED – which are prepared in the course of Task 7.4 (Dissemination and Communication) – will be made available on the greenSPEED website ("Results & Media") in PDF format (approximately 100 MB). Partners can further access the material via SharePoint.

Project videos for dissemination and communication activities prepared within Task 7.4 will be published on YouTube, the greenSPEED website and social media (e.g. LinkedIn). They will also be available via SharePoint to all partners. The videos will be made available in MP4 format (approximately 100 GB, depending on the length of the video).

Dissemination and Communication Activities

Data from dissemination and communication activities will be used for the regular monitoring of these activities within Task 7.4. The monitoring and reporting is only for internal use and is open to the greenSPEED partners during the project. The planned and performed activities such as publications, conferences or social media posts will be monitored via SharePoint (further explained in D7.4 „Preliminary Dissemination, Communication and Exploitation Plan" due in M6 [2]). Partners are required to report any dissemination and communication activities in a SharePoint list. This list can then be exported as xml file (approximately 5 MB, depending on the number of items in the list).

Publications

Open access should be granted according to the provisions of the CA [3] and GA [4] of greenSPEED. All publications should be made available open access (gold or green open access). These publications are also monitored via a SharePoint list which can be exported as xml file (approximately 5 MB, depending on the number of items in the list.). They will further be identified with a DOI (Digital Object Identifier).

5. FAIR Data

As indicated in the previous chapters, data of different nature will be generated during the greenSPEED project implementation. This chapter will provide further inputs regarding the FAIR principle [1].

5.1 Making data findable, including provisions for metadata

The Horizon Europe Programme Guide [1] stresses the importance of data being “FAIR” (Findable, Accessible, Interoperable and Re-usable). All project partners will adhere to common data management principles to ensure that data collected, processed and/or generated during the project is properly managed, archived and retained in accordance with the FAIR principle, even if access is restricted. Data created within greenSPEED will be stored in the SharePoint specifically set up for the project by VIF and an access control list will be established according to the confidentiality levels described in the next subchapter.

Additionally, searching the data will be supported by the classifiers and metadata defined within this chapter. The idea of introducing these classifiers and metadata is to easily search through the datasets, first to provide a smooth interaction and second for synchronisation on data and information between each partner and therefore reducing time consuming communication activities.

All scientific publications will be identified using a DOI and the respective data that has been used for the research will be made available in open-access repositories such as ZENODO [6] or OpenAIRE [7], if possible according to the CA [3] and GA [4]. More details are provided in Chapter 5.3.

In D7.1 “Quality Assurance and Risk Management Plan” [9], the naming conventions for general documents as well as deliverables have been defined. These guidelines are included in the following (see Table 5-1).

Table 5-1: Naming Conventions

| Naming Conventions | |
|--------------------|--|
| General Documents | greenSPEED_Title_yyyymmdd_vX.X.doc/pdf/xls (...) |
| Deliverables | greenSPEED_DX.X_Title_vX.X.doc/pdf (...) |

- **Title:** Short document description (please do not use any special characters such as “.”)
- **Date:** Date of creation (format: yyyymmdd)
- **Version:** vX.Y (X = major version; Y = minor version; such as v0.1, v0.2, v1.0...)
- **File Extension:** According to the type of the file (such as docx, pdf, xls ...)

5.2 Metadata Pre-definition

The metadata of datasets are actively defined at the very beginning of the project to meet the project requirements. A first data management workshop was held, and first metadata and classifiers were defined. More details are provided in the following:

Type of dataset:

- Primary/Secondary data: New generated or existing data (see Chapter 4)

Responsible partner: Describes who is responsible for the data generation and quality

greenSPEED cell generation:

- Gen0/Gen1/Gen2: The datasets are referenced to a generation of the greenSPEED cell generations.

Cell integration level:

- Anode/cathode/cell: This classifier is following the integration process of the cell, from component to full integration of the cell.

Data relation:

- Function/process/material: This classifier provides information on how the data relates to the property of the dataset. A function dataset relates to data describing the function of a cell component or the cell in general. A process classified dataset relates to production process data, most likely describing process related data like energy consumption, all relevant process parameters over time as well as the processing time. Material datasets are based on material data coming from partners, like physical characteristics on single components, e.g. foil properties, properties of active materials.

Dataset name: Describes the properties of the dataset, in a short form

Data format description: Describes the format of the dataset provided

(Expected) amount: Describes the amount of data and allows for preparation on the later data handling

Management system: Provides an answer, where the data is managed and stored

Confidentiality level of datasets:

- Public (PU): Fully open (can be shared with the public)
- Sensitive (SEN): Limited under the conditions of the GA [4] (cannot be shared publicly)

Further additional classification: The greenSPEED project will extend these classifiers during the project runtime. Further information such as due date, work package of origin, supporting partners, whether it is measured or simulated data and more might be provided/added in the future.

5.3 Making data accessible

The greenSPEED SharePoint has been set up during the project proposal phase and has been updated to meet the needs for the project implementation phase. Data can be stored and accessed through SharePoint. The access is restricted to the project partners and can be requested directly from the project coordinator (VIF). As the owner of the greenSPEED SharePoint, VIF can monitor the version history, establish a regular backup policy and restore and manage the stored data.

Publicly available information such as public project deliverables, publications or dissemination material will be accessible via the greenSPEED website [5] and other communication channels (such as LinkedIn). It is not necessary to request access data for the download, as these materials will be made available for the public. In addition, public research data will be publicly available at data platforms like ZENODO [6] or OpenAIRE [7]. The partners are committed to make publications accessible via green open access (self-archiving) or gold open access (open access via the publisher). Either way, partners can choose the option they prefer but must however ensure the open access to the deposited version of the publication.

According to Article 17 “COMMUNICATION, DISSEMINATION, OPEN SCIENCE AND VISIBILITY” and Annex 5 of the Grant Agreement [4], each beneficiary must ensure open access to the scientific publications of its results in a repository for scientific publications and the web-based greenSPEED repository, retaining sufficient IPR to meet with the open access requirements.

5.4 Making data interoperable

The interoperability is achieved by a good implementation of metadata information. Interoperability will be increased if proper metadata standards, methodologies and naming procedures are adopted. At the beginning of this chapter, a first overview of how datasets will be classified is provided. Common data formats such as .docx, .pdf, .mp4, .xls, .txt will be used. Publications must be “machine readable”, which means that scanned versions of publications should not be made available, but text file formats should be chosen. The reader must be able to read it online, download and print it free of charge with no access restrictions.

Usage of commonly used ontologies and a formal and broadly applicable language (e.g. English) will be applied in greenSPEED to enable and improve inter-disciplinary interoperability of (meta) datasets. However, as a living document, improvement identified throughout the project may be introduced in the updated version of the DMP.

It is important that data is stored properly according to the data management guidelines of the greenSPEED project and the FAIR principle [1].

5.5 Increase data re-use

The re-usability factor will be assessed for each dataset in a suitable and easily accessible format. The challenge here is to find a platform that is acceptable to each partner. This is an ongoing work during the first project year. There, the degree of re-usability of the generated data will be determined by the field “confidentiality level”, see also “classifiers” in Chapter 5.2.

Unless otherwise stated, all the data shared and generated in greenSPEED will be considered confidential according to the Consortium Agreement [3] signed by all partners. Nevertheless, considering the collaborative spirit of the project and the academic vocation of some of the beneficiaries involved in the project, re-usable data for dissemination activities will also be generated; the field “confidentiality level” will thereby allow identifying the re-usable data generated during the project.

6. Allocation of Resources

6.1 Costs for making data FAIR

The costs for setting up and maintaining the greenSPEED SharePoint are covered by the coordinator (VIF). After the end of the project, the repository will not be updated. Nevertheless, the files stored there will still be available for access 4 years after the project end.

The costs for publications are covered by the respective partners/authors. Resources and efforts have been allocated to the partner's budgets. If eligible and allocated, costs for open access publications are covered. Further details are outlined in the Grant Agreement [4].

6.2 Responsibilities of data management

As the coordinator, VIF is responsible for the set up and update of the SharePoint, maintenance (structure, upload, downloads etc.), security assessment and respective mitigation measures as well as the access to the SharePoint and the management of user requests. The quality of the data is the primary responsibility of the data provider (the greenSPEED partners). All greenSPEED partners are responsible for handling the data according to the FAIR principle and making it available as openly as possible.

6.3 Long-term preservations

The greenSPEED partners consider the long-term preservations of the data to ensure that this is accomplished. In addition to the SharePoint, also the website will be available 4 years after the end of the project. Afterwards, the website will be archived. The partners will discuss this also with the Advisory Board, if necessary.

7. Data Security, Ethics and Other Issues

7.1 Data Security

The greenSPEED consortium makes every effort to protect the data, products and services from unauthorised use and provide secure access to data. All shared and processed greenSPEED data will be stored in a secure environment with access rights restricted to the respective project partners. If data is to be transferred between the partners, the respective articles in the CA [3] and GA [4] must be complied with.

7.2 Ethics

greenSPEED will be carried out with the highest ethical standards and the applicable EU, international and national law on ethical principles. Ethics and legal issues are covered in the CA [3] and GA [4] of the greenSPEED project. During the ethical review performed by the funding authority, no ethical issues have been identified and the project has been cleared. The project consortium regularly evaluates possible ethical issues that may arise during the project implementation. Currently, no ethical issues have been identified by M6. Further ethical issues will be added here as needed.

7.3 Other Issues

Currently, no other issues have been identified. Further issues will be added here as needed.

8. Conclusion

In a first step, the initial data management plan for greenSPEED was defined together with the consortium. The partners are aware of the regulations and recommendations of the European Commission and are willing to support these actions. In the first six months, a first data management workshop was already held with the partners. There will also be workshops on this topic during the project, whenever necessary.

The greenSPEED partners will adhere to the FAIR principle [1] with respect to the IPR and regulations set in the Consortium Agreement [3] and Grant Agreement [4]. It is essential to note that this Initial Data Management Plan (D7.2) is a first draft and will be evaluated and updated on a regular basis. This is important to ensure a high-quality data management process and the sustainability of the greenSPEED results.

9. Abbreviations

| Term | Definition |
|------------------|--|
| CA | Consortium Agreement |
| DM | Data Manager |
| DMP | Data Management Plan |
| DOI | Digital Object Identifier |
| D(7.2) | Deliverable (7.2) |
| FAIR (data) | findable, accessible, interoperable and re-usable (data) |
| GA | Grant Agreement |
| GDPR | General Data Protection Regulation |
| Gen | Generation |
| greenSPEED | Green and Sustainable Processes for Electrode Production |
| M6 | Month 6 |
| MW-PECVD | Microwave-Assisted Plasma Enhanced Chemical Vapor Deposition |
| Ni-rich NMC | Nickel-Mangan-Cobalt cathode material with high Ni content |
| PU | Public |
| R | Document, Report |
| SEN | Sensitive |
| SiH ₄ | Silane gas |
| TRL | Technology Readiness Level |
| VOCs | Volatile Organic Compounds |
| WP | Work Package |

10. References

- [1] Horizon Europe Programme Guide: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf, Version 2.0, 2022-04-11 (last accessed December 2022)
- [2] greenSPEED Deliverable D7.4 „Preliminary Communication, Dissemination and Exploitation Plan”, planned submission: December 2022
- [3] Consortium Agreement greenSPEED, Version [final 7.0], 2022-08-01
- [4] Grant Agreement Number 101069528 – greenSPEED, 2022-06-15
- [5] www.greenspeed-project.eu
- [6] www.zenodo.org
- [7] www.openaire.eu
- [8] Horizon Europe Data Management Plan Template, Version 1.0, 2021-05-01
- [9] greenSPEED Deliverable D7.1 “Quality Assurance and Risk Management Plan“ v2.0, 2022-10-28