



Flexible energy systems Leveraging the Optimal
integration of EVs deployment Wave

Grant Agreement N°: 101056730

Deliverable 9.2

Data & Knowledge Management Plan

Authors: Cristina Corchero, Anzhelika Ivanova, Sarah Cnockaert, Josh Eichman
[IREC]



Funded by
the European Union

[Website FLOW](#)

Document Information Table

Project Data			
Project Acronym	FLOW		
Project Title	Flexible energy system Leveraging the Optimal integration of EVs deployment Wave		
Grant Agreement n.	101056730		
Topic identifier	HORIZON-CL5-2021-D5-01-03		
Funding Scheme	RIA		
Project duration	48 months		
Coordinator	Catalonia Institute for Energy Research		
Website	https://www.theflowproject.eu		
Deliverable Document Sheet			
Deliverable No.	9.2		
Deliverable title	Data & Knowledge Management Plan		
Description	The DMP intends to address the lifecycle and public availability of research data generated by the project; and lays down the framework for governing issues related to data management.		
WP No.	9		
Related task	T9.1, T9.2		
Lead beneficiary	IREC		
Author(s)	Cristina Corchero, Anzhelika Ivanova, Sarah Cnockaert (IREC)		
Contributor(s)	Josh Eichman [IREC]		
Type	R — Document, report		
Dissemination L.	PU - Public		
Due Date	31/12/2022	Submission Date	03/01/2023

Version	Date	Author(s)	Organisation	Comments
V0.1	15/11/2022	Cristina Corchero, Sarah Cnockaert	IREC	Deliverable structure and contents
V0.2	25/12/2022	Cristina Corchero, Anzhelika Ivanova	IREC	First draft
V1.0	03/01/2023	Cristina Corchero, Anzhelika Ivanova Sarah Cnockaert	IREC	Version for submission

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 document and all information contained here is in the sole property of the FLOW Consortium. It may contain information subject to Intellectual Property Rights. No Intellectual Property Rights are granted by the delivery of this document or the disclosure of its content. Reproduction or circulation of this document to any third party is prohibited without the written consent of the author(s). The dissemination and confidentiality rules as defined in the Consortium agreement apply to this document. All rights reserved.

Table of Contents

Document Information Table	2
DISCLAIMER	4
Table of contents	5
List of Acronyms	6
Executive Summary	7
1. Data Summary	8
1.1. Data collection/generation	8
1.2. Re-use of existing data	8
1.3. Types and formats of data	8
1.4. Origin of data	8
1.5. Expected data size	9
1.6. Data utility	10
2. FAIR data	11
2.1. Making data findable, including provisions for metadata	11
2.2. Making data accessible	11
Repository	11
Data	12
Metadata	12
2.3. Making data interoperable	13
2.4. Increase data re-use	13
3. Other research outputs	14
4. Allocation of resources	14
5. Data security	15
6. Ethics	15
7. Other issues	15

List of Acronyms

ACRONYM	MEANING
DMP	Data Management Plan
DoA	Description of Action
EV	Electric Vehicle
EU	European Union
WP	Work Package

Executive Summary

The Horizon Europe Model Grant Agreement requires that a Data Management Plan (DMP) is established and updated regularly. The DMP intends to address the lifecycle and public availability of research data generated by the project; and lays down the framework for governing issues related to data management. This includes procedures on how to work with the data within the lifecycle of the project and a description of datasets. The DMP is intended to be a living document in which information can be made available on a finer level through updates as the implementation of the project progresses and when changes requiring updating the DMP occur.

FLOW's DMP is based on the recommended template for Horizon Europe beneficiaries. In completing the sections of the template, the requirements for research data management of Horizon Europe as described in article 17 and analysed in the Annotated Grant Agreement, article 17, are addressed.

1. Data Summary

This section provides information on the data that will be used and generated during the project. It is worth noting that due to the early stage of the project, specific details may be missing (indicated throughout the document). These will be discussed in a workshop at the next General Assembly, 30-31st of January 2023, and updated in future versions of this deliverable.

1.1. Data collection/generation

What is the purpose of the data generation and its relation to the objectives of the project?

The main purpose of the data generation is to provide the necessary inputs for the creation of the software tools stated in the Description of Action (DoA). The software tools will also generate output data.

1.2. Re-use of existing data

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

Although the specific data sets have not been defined yet, in order to develop the new solutions and tools the use and re-use of the following existing data sets can be expected:

- Vehicle driving behaviours (no specific sources yet)
- Vehicle charging behaviours if available (individual vehicles)
- Battery degradation measurements (public data sources, e.g., battery archive, other EU projects, the literature)
- Single facility electricity consumption
- Weather data
- System level data (e.g., power capacity by generator, transmission capacity, energy demand)
- TEA data (e.g., capital cost, operation and maintenance, efficiency, equipment lifetime)
- Data from individual equipment characterization

1.3. Types and formats of data

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

The types and formats of the data depend on whether it is input or output data. Input data will follow the formats already established by the partners/other sources that will provide this data. Output data formats depend on the tools that will be defined and developed during the project. As FLOW is in the beginning stages, the exact types and formats have not been determined yet and will be updated as the project progresses.

1.4. Origin of data

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

The input data, such as the vehicle driving and charging behaviour, the equipment and facilities data and the system level data, will come from the corresponding partners of the project. Other input data, such as weather and battery degradation data, will come from outside resources (public data sources, e.g., battery archive, other EU projects, the literature) that will be determined further on.

1.5. Expected data size

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

The following data sets can be expected for every WP:

- WP2
 - Survey data is less structured
- WP3
 - Data space and data exchange for use cases
- WP4
 - Input
 - Vehicle driving behaviors (no specific sources yet)
 - Vehicle charging behaviors if available (individual vehicles)
 - Battery degradation measurements (public data sources, e.g., battery archive, other EU projects, the literature)
 - Single facility electricity consumption
 - Weather data
 - Output
 - Vehicle charging behaviors (timeseries)
 - Model of battery degradation
 - Forecasted electricity demand for facilities (timeseries)
 - Optimized operation profiles for bidirectional charging of EVs, stationary storage, flexible demand (structured timeseries data).
- WP5
 - System level data (e.g., power capacity by generator, transmission capacity, energy demand)
 - Vehicle charging behaviors (for entire region)
 - TEA data (e.g., capital cost, operation and maintenance, efficiency, equipment lifetime)
- WP6

- Data from individual equipment characterization
- Data generated in testbeds
- WP7
 - Data generated in demos

The size of the data sets cannot be determined at this stage in the project as the final data sets have not been determined yet. However, an estimation can be given. If we assume approximately more than 1 GB for every data set in the list given above, the size of data in the project can be expected to be higher than 17 GB.

1.6. Data utility

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

Outside of the project, the data from FLOW can be used by EV vehicle manufacturers and users, electrical utilities, battery manufacturers, as well as other Horizon Europe projects. The extent of the data that will be publicly available will depend on certain restrictions. Some of the data sets will not be available for open use if they are intended for further exploitation or if they pose any security and confidentiality issues.

2. FAIR Data

This section describes the measures taken to facilitate identification, interoperability, access and sharing of data within the project and with the rest of stakeholders. In this regard, the key objective of the FAIR DMP state the requirement of making research data Findable, Accessible, Interoperable and Re-usable (FAIR) aiming to making it useful for future research and developments.

2.1. Making data findable, including provisions for metadata

- *Will data be identified by a persistent identifier?*

A persistent identifier will be necessary in order to provide a long-lasting digital reference for each open data set. The most commonly used identifier is the Digital Object Identifier (DOI), but the choice of the type of persistent identifier will depend on the data repository that will be chosen in the next stages of the project.

- *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.*

This will be determined following the list of metadata standard formats. An example is available at <https://www.dcc.ac.uk/guidance/standards/metadata/list> .

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

Yes, keywords will be available for each open data set to optimize the possibility for discovery and re-use.

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

The intention is to provide metadata in this way.

2.2. Making data accessible

Repository

- *Will the data be deposited in a trusted repository?*

The data will be deposited in repositories that are trustworthy and suitable for the type of deposited data. The repositories will be identified in later versions of this deliverable, when the types of data are known in more detail.

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

The arrangements will be done after the identification of the repositories.

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

Most of the trustworthy repositories assign an identifier to the data. This will be considered as one of the criteria when choosing the repository.

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.*

Not all data will be openly available due to confidentiality, security and exploitation reasons. The restrictions on the data sets will be explained in detail once the exact data sets are determined.

- *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.*

If any embargo is applied, it will be done in the later stages of the project, when the tool architecture is defined, as well as the input and output data.

- *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)?*

These aspects will be determined in a workshop on the next General Assembly, 30-31st of January 2023.

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?*

When and if possible, the aim is to make metadata openly available and licenced under a public domain dedication CC0. In the cases where this will not be possible, a detailed justification will be provided.

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

These aspects will be determined in a workshop on the next General Assembly, 30-31st of January 2023.

- *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)?*

Documentation is can be provided with the open software. Whether this will be needed to read or access the data will be known once the different software architectures are finalized.

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?*

The aim is to follow community-endorsed interoperability practices to facilitate data exchange and re-use. The exact practices will be defined among partners depending on the data in the further stages of the project.

- *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)?*

These aspects will be determined in a workshop on the next General Assembly, 30-31st of January 2023.

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.)?*

These aspects will be determined in a workshop on the next General Assembly, 30-31st of January 2023.

- *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?*

In all possible cases, data will be made freely available in the public domain, licensed following the standard reuse licences as stated in the Grant Agreement. In the cases where this will not be possible, it will be duly justified.

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

These aspects will be determined in a workshop on the next General Assembly, 30-31st of January 2023.

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

Yes, the data will be documented following the appropriate standards.

- *Describe all relevant data quality assurance processes.*

There are many ways to assure data quality. The exact process for data quality assurance will be determined in the next stages having into account the following principles:

- Relevance – the provided data can be interpreted by the software and legal conditions allow its use
 - Accuracy – implement data filtering to ensure accuracy of the data and remove bad data
 - Consistency – internal and external check to assure consistency of the data
 - Timeliness – up to date data
 - Compliance – ensure that the data complies with legal and contractual obligations
- *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.*

This will be taken into account and these aspects will be addressed in detail in the next version of the DMP.

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.*

This will be discussed in the next General Assembly, 30-31st of January 2023, and addressed in detail in the next versions of the DMP.

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).*

This will be discussed in the next General Assembly, 30-31st of January 2023, and addressed in detail in the next versions of the DMP.

- *Who will be responsible for data management in your project?*

As responsible for the DMP, IREC will follow closely the data management process. However, the responsible for the data management will be determined in the General Assembly, 30-31st of January 2023.

- *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)?*

This will be discussed in the next General Assembly, 30-31st of January 2023, and addressed in detail in the next versions of the DMP.

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)?*

This will be discussed in the next General Assembly, 30-31st of January 2023, and addressed in detail in the next versions of the DMP.

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

This will be done for all data sets where it is possible and/or relevant.

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 DoA.*

This will be discussed in the next General Assembly, 30-31st of January 2023, and addressed in detail in the next versions of the DMP.

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

Yes, informed consent will be included if there are cases with data sharing and long term preservation of 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)?*

This will be discussed in the next General Assembly, 30-31st of January 2023, and addressed in detail in the next versions of the DMP.