D1.2 – Data and Knowledge Management Plan

SWITCHING-CELL-ARRAY-BASED POWER ELECTRONICS CONVERSION FOR FUTURE ELECTRIC VEHICLES

DATE: 30 December 2022   **VERSION:** 2.0

Author(s): Àlber Filbà (IREC), Gabrielle Lacube (IREC)

Contributor(s): Josep Maria Herrera (IREC), Luis A. López (IREC), Sergi Busquets (UPC), Mehmet Usta (AVL)

Dissemination level: Public

Work package: WP1

Description: Definition of the items that are subject to intellectual property and the procedures to exchange information among the partners. Updates will be informed during project reports.

Executive Summary

The Data and Knowledge Management Plan (DKMP) intends to address the lifecycle and public availability of research data generated by the project, laying down the framework for data management. This includes procedures on how to make the data findable, accessible, interoperable, and reusable, the so-called FAIR principle. The DKMP also proposes a preliminary definition of data types, formats, ownership of data and data driven research results, as well as Intellectual Property Rights, access, data sharing, storage, and preservation.

The DKMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when changes requiring updating the DMP occur.

Document History

Date	Person	Action	Status
9 December 2022	Àlber Filbà (IREC) Gabrielle Lacube (IREC)	First version	Draft (V0.1)
12 December 2022	Luis Alberto López	Revision of data types	Revision (V1.0)
14 December 2022	Mehmet Usta (AVL)	Revision of IPR statements	Revision (V1.1)
16 December 2022	Josep Maria Herrera (IREC)	Minor corrections	Revision (V1.2)
20 December 2022	Sergi Busquets (UPC)	Minor corrections	Revision (V1.3)
30 December 2022	Àlber Filbà (IREC)	Consolidation of all revisions	Final (V2.0)

List of Figures

List of Tables

Contents

1. Introduction

This document is the SCAPE project Data and Knowledge Management Plan (DKMP). It describes the data management life cycle for all datasets that will be collected, processed, and generated by this Research and Innovation Action project. The DKMP will be updated during the project lifetime.

The present document covers:

- Intellectual Property Rights (IPR) management
- The management of research data during and after the project lifetime.
- Nature of data to be collected and processed.
- Provisions to comply with FAIR data standards.
- Data storage issues.

2. IPR Management

When considering IPR in this project, it is important to emphasize the high degree of innovation and exploitability of the proposed power electronics design methodologies, integration technologies, control strategies, and testing techniques. In this respect, the partners in the Consortium acknowledge the strategic importance and sensitive nature of data provided by any other partner, and that disclosure of such information will be subject to specific obligations with respect to confidentiality and use.

IPR issues are governed by the provisions of the Grant Agreement and Consortium Agreement. Particularly, any background, particular ideas, or intellectual property developed by a single partner will remain the property of the holder. Shared IPR between partners must be established in writing in a "joint ownership agreement". Additionally, the consortium partners must grant each other access to the background identified in the Consortium Agreement and the results developed in the project for the correct implementation of the project activities and the exploitation of the results. The access must be made on a royalty-free basis and under fair and reasonable conditions. Requests to exercise access rights and the waiver of access rights must be in writing.

3. Data Summary

This section details the data types generated and used in the project. More in-depth details will be provided in future updates of the DKMP, released as the implementation of the project progresses.

The purpose of data collection and generation within the SCAPE project is:

- To measure the results of project experiments.
- To compare and share the experiments results among the partners, stakeholders, and prospective audience.
- To perform the life-cycle assessment and compute the life-cycle cost of the proposed solution and compare with the traditional one.
- To assess the impact of the project outcomes.
- To disseminate the project results.

For this matter, the project coordinator has created a repository for the project through the Zenodo Digital Library for all the project partners to make publicly available any project open-access data. The SCAPE Zenodo repository (also named "SCAPE community") is accessible through link: https://zenodo.org/communities/scape-power-he. Figure 1 shows its frontpage. Section 4.2 provides more detail on how the repository works."SCAPE Project Data Repository".
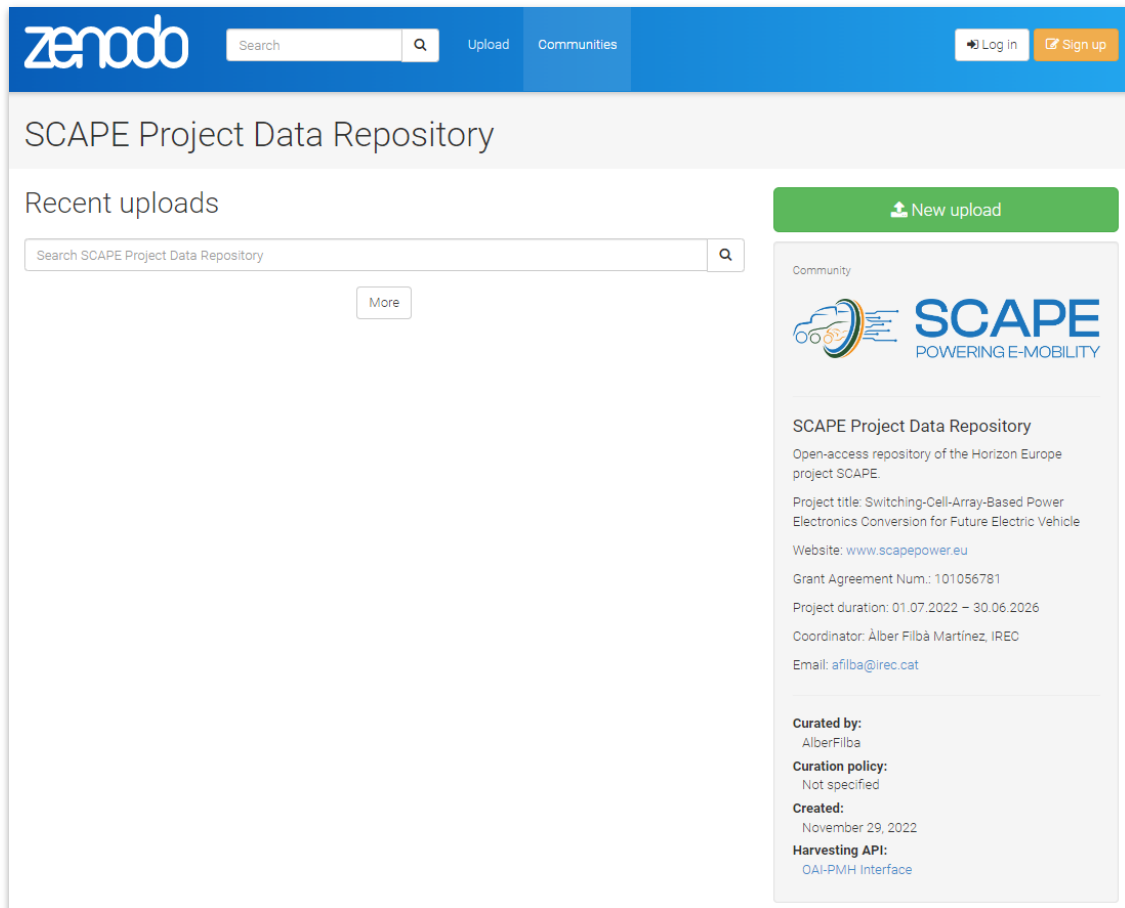

Figure 1. SCAPE Zenodo community frontpage. Uploaded datasets, documents and other media will appear here.

3.1. Types and formats of data

The data generated and used in the project is related to the project technical objectives and to relevant communication, dissemination, and exploitation activities. Thus, in this context we understand "data" in a broad sense, including not only raw information obtained from the design, simulation, and experimental activities, but also multimedia data and documents aimed to the communication and dissemination of the project outcomes.

Table 1 summarizes the main data-item types and its file formats, distinguishing between the "internal-use file format" meant for data shared between the consortium partners, and the "open-access file format" meant for data made publicly available. The open-access data formats allow using open-source data-treatment software, committing to the FAIR data standards (see Section 4), enabling the widespread use of the data, and easing the access to the dissemination material.

The data will be generated within the technical activities and will be complemented by existing data provided by the partners, the community, and/or from literature. Use of non-original data will be done only upon authorization from the original authors, acknowledging its authorship in case the data is made publicly available; e.g., under Creative Commons licensing. The experimental data is considered long-term value and should be preserved as state-of-the-art data at the time of data production.

Table 1 – Data-item types generated and used in the project.

Data item type	Internal-use file format	Open-access file format	File size
Experimental test results datasets	Platform specific	CSV	Up to tens of gigabytes
Simulation datasets	Platform specific	CSV or plain text file (TXT)	Up to tens of gigabytes
Multiphysics simulation models	Platform specific	Formats from open-source (OS) software tools (e.g., Elmer software)	Up to units of gigabytes
Mechanical diagrams	Platform specific	PDF, 3D-PDF, STEP, or STL	Up to hundreds of megabytes
Electrical schematics	Platform specific	Plain text file (netlist) or PDF	Up to tens of megabytes
Printed-circuit board designs	Platform specific	Gerber or PDF	Up to hundreds of megabytes
Software code	Text file	Plain text file (TXT), OS word-processor file, or PDF	Up to tens of megabytes
Survey results datasets	Platform specific	Plain text file (TXT), OS word-processor file, PDF, or CSV	Up to tens of megabytes
Technical-scientific documents	Word-processor file	OS Word-processor file or PDF	Up to tens of megabytes
Life Cycle Inventory questionnaires	Spreadsheet file	CSV or PDF	Up to tens of gigabytes
Technical-environmental datasets	Spreadsheet file and word-processor file	CSV or PDF	Up to tens of gigabytes
Life Cycle Analysis results	Platform specific (GaBi) and spreadsheet file	CSV or PDF	Up to tens of gigabytes
Infographic documents	Platform specific	PDF, SVG, PNG, or TIFF	Up to tens of megabytes
Images	JPG, PNG, or TIFF	JPG, PNG, or TIFF	Up to tens of megabytes
Videos	MP4, AVI, MOV…	H.264-encoded MP4	Up to tens of gigabytes

3.2. Data security

The project coordinator (IREC) has created a cloud file server based on Microsoft SharePoint to share files and enable online collaboration between the consortium members. The project SharePoint has a file version tracking, allowing automatic backup of files. Nonetheless, if a file is deleted to the SharePoint recycle bin, it is not possible to recover it once the recycle bin is emptied or 90 days have passed after the file deletion. Thus, it is advised that each partner is in charge of backing up their data locally.

Additionally, the project coordinator has created the Zenodo community "SCAPE Project Data Repository" for all the project partners to make publicly available any open-access data (see Section 4.2). Zenodo will also serve as a permanent repository of the project outcomes after the project has ended.

3.3. Data utility

All data generated is expected to support project activities and tasks. Additionally, the project data and information may be of interest and may be used by different sets of stakeholders including the advisory board members, industrial stakeholders, and academia.

4. FAIR Data

This section describes the required measures 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 Data Management plan is ensuring that research data is Findable, Accessible, Interoperable, and Re-usable (FAIR), making it useful for future scientific and technological research, innovation, and development activities.

4.1. Making data findable, including provisions for metadata

To make data findable, it must be ensured that the project open-access data can be discoverable thanks to its associated metadata. When uploading a data item to the SCAPE Zenodo repository, the user must compulsorily fill in several fields to provide the necessary metadata for the proper description of the data item. Figure 2 shows the most relevant fields in the Zenodo data-uploading site[1]. Among them, the following fields are of utmost importance:

- Data Type (aka Upload Type)
- Digital Object Identifier (DOI)
- Title
- Authors
- Version
- Access Rights
- Keywords
- Grants

---

[1] https://zenodo.org/deposit/new?c=scape-power-he

Figure 2. Fields for the user-provided metadata appearing in the data item upload site of the SCAPE Zenodo community.

The DOI metadata is assigned to each data item so they can be easily findable and discoverable, through persistent and unique identification. Zenodo automatically assigns a unique DOI (e.g., 10.5281/zenodo.7418756) to any uploaded data. The data-item metadata can be modified without requiring a new DOI. Nonetheless, updating the data-item main content forces increasing its version number, generating a new DOI.

Data-items title naming must obey the following naming convention:

- A prefix "SCAPE" indicating that is a dataset from the project.
- Type of the data ("dataset", "publication", "video", etc.), followed by " – ".
- Title of the dataset.
- Version number.

Version numbering follows a "major.minor" rule, similar to software versioning systems (e.g. v2.1), in a different manner than Zenodo automatic number, where numbering is linear. An example of data-item name is "SCAPE Dataset - Title v1.0". The defined title naming may be also used for the data files naming.

To maximize the possibilities of re-use of datasets, keywords must be provided to the data items. Basic keywords applicable to all datasets may be "Horizon Europe", "SCAPE project", or "Electric Vehicle", while more specific keywords may be for example "power converter", "thermal model", or "powertrain architecture". Additional technical keywords specific to SCAPE open datasets will be identified and provided at a later stage of the project.

Access right should be in most cases "Open Access" with a Creative Commons license (see Section 4.4 for further detail). "Embargoed Access" should be employed for journal/conference publications following the "green" open-access route[2], when delayed open access is required by the journal/proceedings policy. "Restricted Access" and "Closed Access" may be used for special cases when consortium internal data sharing is required, although the SCAPE SharePoint is preferred.

Any data item uploaded to the SCAPE Zenodo community must be provided with the SCAPE project grant number (101056781) in the "Grant" metadata. If the data item is related to any other research projects, these can be added after the SCAPE grant metadata entry.

All the generated metadata can be readily harvested and indexed through Zenodo harvesting utility[3].

4.2. Making data accessible

The research data generated in SCAPE Project not subject to IPR will be made publicly available. This is of special relevance for the datasets obtained and/or used to publish results, in order to ensure replicability of such results. A data repository is necessary to facilitate deposit, update, management, and reference of the data generated and used within the project. Since not all partners have access to an institutional repository, the consortium, as already introduced in Section 3.2, has selected Zenodo since it ensures unified management procedures for any kind of the data-item types listed in Table 1.

---

[2] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

[3] https://zenodo.org/oai2d?verb=ListRecords&set=user-scape-power-he&metadataPrefix=oai_dc

Zenodo is an all-purpose open research repository. It was created by CERN and OpenAIRE, a non-profit partnership to support European research, to provide a place for researchers to deposit publications, datasets and other research artifacts such as code, posters, presentations, etc. Zenodo does not impose any requirements on format, size, access restrictions or license, and is not restricted to one funder or one nation. It is free to use for all research outputs from across all fields.

Thus, the SCAPE Zenodo community[4] has been created and it will be linked to the project website to foster its use among the general public. Partners are required to deposit any data item listed in Table 1 to the community.

Despite the closed access solutions offered by Zenodo, the Consortium will make use of SCAPE SharePoint internal repository to share among consortium partners and relevant project stakeholders; e.g., the Advisory Board members, any data subject to confidentiality or IPR.

4.3. Making data interoperable

Making data interoperable means that the project consortium must ensure data exchange and re-use between researchers, institutions, organizations, countries, etc. Its main goal is facilitating the re-combination of data items from other origins, especially when concerning datasets, models, and software code. To achieve this goal, the consortium must promote the use of standard formats and of open-source software applications within the project activities. Additionally, the consortium must attempt collecting and documenting the data items in a standardized way to ensure that these can be correctly understood, interpreted, and re-used.

In SCAPE, data items should be documented with a description of their main variables, the general methodologies employed for the generation of the data, and a general glossary with the main employed vocabulary. On the latter, standard vocabulary will be used to allow inter-disciplinary interoperability.

4.4. Increase data re-use

All data items obtained within the project and deemed accessible will be made available for re-use. To avoid any potential doubt, the consortium will provide as part of the data-item metadata the deposited data Creative Commons (CC) license, defining all conditions under which the work is provided.

CC Licenses contain four license features: BY, NC, ND, and SA. The combination of these features defines six possible CC License variations, as shown in Figure 3. The description of the license features are:

- "BY" stands for Attribution: The obligation to credit the author and other parties designated for attribution.
- "NC" stands for NonCommercial: Commercial use is excluded from the license grant
- "ND" means NoDerivatives: Only verbatim copies of the work can be shared
- "SA" represents ShareAlike: The work can be modified, and modified versions can be published but only under the original or a compatible license.

Zenodo offers multiple variations and versions of the CC Licenses, all including the attribution feature (CC BY) to appropriately credit the authors for the original creation. When possible, the

---

project datasets license must be "CC Attribution 4.0 International (CC BY 4.0)" allowing third parties to share and adapt data with no restrictions as long as attribution is provided. Any of the other six variations of the CC Licenses may be used, but version 4.0 International must be used in any case.

All data will be stored in Zenodo as soon as possible and no later than the publication of the related scientific publication and will remain re-usable for the lifetime of the repository, which is currently warrantied for a minimum of 20 years.

Currently, there is no specific data quality-check process foreseen within the project. All datasets that will be made available will have been used for the project research purposes and thus any potential error, inaccuracy, or omission in the dataset shall be detected prior to its publication. However, if this becomes relevant, a data quality assurance process will be set.
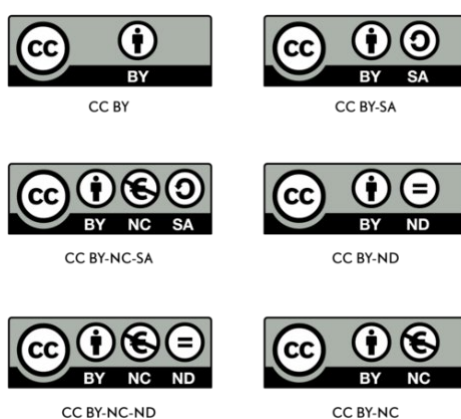

Figure 3. Banners representing the six variations of the CC licenses, by Markus Büsges, licensed under CC BY 4.0.

5. Allocation of resources

IREC as project coordinator will be key responsible for DKMP compliance as well as the FAIR data strategy implementation. In this regard, all partners will be requested to ensure the proper upload of data into Zenodo and SharePoint repositories. In this line, the project is fully committed with open access publication, promoting the publication of the project research articles through the "gold" open-access route, with budget already allocated to this. Nonetheless, in cases where publication to open-access journals is not possible, consortium partners should follow the "green" open-access publication route (self-archiving), publishing a preprint version of the article to SCAPE Zenodo open repository.

It is worth noting that all project partners have dedicated resources, including working time and publications fees, allocated in their own budget. When considering long term preservation of data, it is important to note that Zenodo offers free data archiving; however, when considering restricted or confidential datasets, own partners repositories are expected to be used as long-term storage. Under justified circumstances, the SCAPE SharePoint repository may be extended for a certain time after the project ends. In any case, both options will be covered with partners' own resources after the project ends.

The data management will be handled by the Executive Board, who will formally ensure that it is done according to the project rules and sharing requisites. The Executive Board is free to determine the format of the presentation and all other aspects of data disclosure, if required.

Finally, it must be indicated that records of the work (including datasets) must be kept by the partners for at least 5 years after project finalization.

6. Ethics

No specific ethics requirements are needed for the research within SCAPE project.

The consortium ensures that all Horizon Europe ethical standards and guidelines will be applied and actively followed, regardless of the country in which the research will be carried out. Additionally, this deliverable states that the research performed outside the EU could have been and can be legally conducted in one of the EU Member States, and is compatible with the Union, National and International legislation.