Grant agreement no:	Project Acronym:	Project title:
711792	**LiRichFCC**	**Novel Cathode Materials for Lithium Ion Batteries**

Organization name of lead contractor:	Document version:
Karlsruhe Institute of Technology	**2017-03-31**

Work package:
WP 5: Management & Evaluation

Title:
D5.3: Data Management Plan

Start date of project:	Contractual Delivery Date:	Actual Delivery Date:
Oct. 1st 2016 **(duration 36 months)**	**2017-03-31**	**2017-03-31**

Dissemination level:
☒ **PU (Public)**
☐ **CO (Confidential, only members of the consortium + Commission)**

Authors (beneficiary name):
Christian Punckt (KIT)
Maximilian Fichtner (KIT)
Robert Dominko (NIC)
Jean-Francois Colin (CEA)

1 EXECUTIVE SUMMARY

The Data Management Plan (DMP) of LiRichFCC gives an overview of research data anticipated to be produced by the consortium, data management within LiRichFCC, assessment of data, and activities planned to make data FAIR.[1] The DMP reflects the current state of the discussions, plans and ambitions of the LiRichFCC partners, and will be updated periodically as work progresses. It is based on the H2020 Guidelines on FAIR Data Management in Horizon 2020 (26 July 2016). **Key audience of this first version of the report is the LiRichFCC consortium** with the aim to **raise awareness of FAIR data management and Open Access**. Future versions will more strongly focus on actual Data Management procedures and policies employed by the LiRichFCC partners, as the rate of data exchange has further increased.

2 INTRODUCTION

As stated in the H2020 Guidelines on FAIR Data Management in Horizon 2020 (26 July 2016), »Data Management Plans (DMPs) are a key element of good data management. A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).«

LiRichFCC will follow procedures and implement measures to make data (as well as associated metadata) which are generated and used in the course of the project accessible and usable not only for the members of the LiRichFCC consortium but also for stakeholders outside LiRichFCC as well as for the general public.

However, as reflected by the fact that more than half of the Deliverables within LiRichFCC are designated as confidential, a significant amount of data generated by the project cannot be publicly shared. In those cases, data management will still be an important issue, as the success of LiRichFCC depends strongly on the efficient exchange of data and flow of information within the consortium.

2.1 Why is data management important?

As laid out on the KIT Guidelines for Responsible and Sustainable Research Data Management, "responsible handling of research data is indispensable for reproducing scientific findings and making them accessible and usable for science and the society. Protection, storage, and sustainable supply of research data require application of specific standards and compliance with legal framework conditions, with particular attention being paid to the observation of data protection regulations and of the general right of privacy. Karlsruhe Institute of Technology (KIT)" and all other LiRichFCC partners support "its scientists in implementing the guidelines for research data management to safeguard good scientific practice and in using existing services and infrastructure facilities for this purpose."

---

[1] FAIR: findable, accessible, interoperable, and reusable

2.2 Why Open Access?

The European Commission supports the Open Access philosophy and states in its H2020 Program Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020 (Version 3.2): "Modern research builds on extensive scientific dialogue and advances by improving earlier work. The Europe 2020 strategy for a smart, sustainable and inclusive economy underlines the central role of knowledge and innovation in generating growth. Broader access to scientific publications and data therefore helps to:

- build on previous research results (improved quality of results)
- encourage collaboration and avoid duplication of effort (greater efficiency)
- speed up innovation (faster progress to market means faster growth)
- involve citizens and society (improved transparency of the scientific process).

This is why the EU wants to improve access to scientific information and to boost the benefits of public investment in research funded under Horizon 2020. The Commission considers that there should be no need to pay for information funded from the public purse each time it is accessed or used. Moreover, it should benefit European businesses and the public to the full. This means making publicly-funded scientific information available online, at no extra cost, to European researchers, innovative industries and the public, while ensuring that it is preserved in the long term. Under Horizon 2020, the legal basis for open access is laid down in the Framework Programme and its Rules for Participation."

2.3 What can be expected from LiRichFCC DMP?

The DMP reflects the current status of discussion within the consortium about the data that will be produced. It is not a fixed document, but evolves during the lifespan of the project.

3 DATA MANAGEMENT IN LIRICHFCC

Research data management starts with the planning of research projects and includes the acquisition, processing, and storage of research data. It guarantees access or access restrictions, reproducibility, and reuse of research data. The researchers shall be responsible for research data and compliance with specific standards.
LiRichFCC will take advantage of advice provided by central service units at the partner institutions relating to research data management in research projects, from planning to execution to beyond the completion of the project. Advice shall consider discipline-specific requirements, such as the selection of suited repositories and data formats. A further important platform to obtain information and organize data management is the platform OpenAIRE which the consortium plans to use for its data management and sharing activities as well.

3.1 What kind of data will be obtained, what format will it have, and what is its utility?

Data generated in LiRichFCC include:

- Parameters and results of theoretical modelling (in various, often proprietary formats)
- Material synthesis protocols (typically in the form of documents/reports)

- Material characterization data (typically in proprietary format)
- Images, schematics, and graphs (in common data formats such as bitmap or jpg)
- Results summary presentations (PowerPoint files)
- Journal articles, patents, reports on project deliverables (PDF documents)

Data will thus be originated by theoretical and experimental research and development activities, certain types of data will be frequently re-used (e.g. for comparison with modified simulation parameters, synthesis protocols etc.), and it will have moderate size on the order of typically few tens of GB.

Depending on the Work Package involved in data generation, data may not only be useful for members inside the consortium but also for other academic institutions or for industry that might want to do benchmarking of new models, protocols or materials in comparison with existing battery technology.

3.2 How will data be managed internally?

All LiRichFCC partners provide appropriate storage facilities for research data and provide controlled accesses as well as appropriate infrastructure. They also support free access to research data considering ethical, legal, economic, and contractual framework conditions.

3.3 What data can be made public?

Experimental data and synthesis protocols won't be made openly available as default, as their results may have the potentiality to be patented. Data contained in journal articles may be made openly available. Concerning deliverables, their confidential or public character is already defined and available on the European H2020 portal.

Some data may be openly available with some delay due to possible patent applications.

For those data that can be made public, it needs to be ensured that it is findable, accessible, interoperable, and reusable (FAIR). To this end, proprietary formats (see 3.1) will be converted into international standard formats such as ASCII and stored as text files. That way, scientists and development engineers from all over the world which are researching on the field of Li-ion batteries or the synthesis and electrochemistry of new Li-rich cathode materials for Li-ion batteries will benefit from the LiRichFCC program.

3.4 What processes will be implemented?

The partners of the LiRichFCC consortium combine over a century of experience in research data handling, and have developed efficient ways to archive and share data. Nonetheless, research has become increasingly more interdisciplinary, and amounts of data generated are on the rise. Therefore, especially for collaborative work within individual work packages, the partners follow internal codes and standards for making data findable.

Parameter sets, methods and protocols will be stored in text documents that follow standardized naming conventions jointly defined by the LiRichFCC partners to ensure maximum findability, accessibility and interoperability.
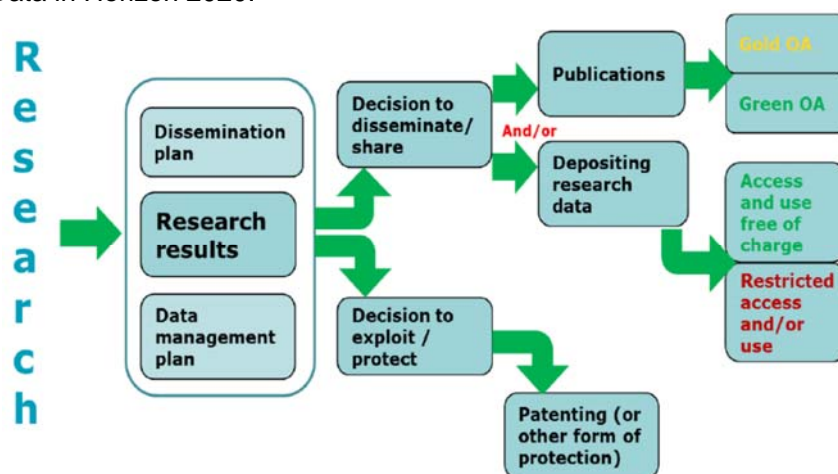
Aggregated data in the form of presentations, reports (deliverables), publications, or patents follow standardized naming conventions. For example, presentations and reports include the name of the project, the corresponding work package, and the date. Deliverables can be identified by their deliverable number, publications have unique DOIs, and patents are numbered per international standards. Public aggregated data will by default be made available on the project webpage (www.lirichfcc.eu) as well as in a yet-to-be-determined professional repository.

Aggregated data to be shared will always be in a format that can be used and understood by a computer. They will typically be stored in PDF formats that are either standardised or otherwise publicly known so that anyone can develop new tools for working with the documents.

Raw experimental or theoretical data that has been identified as non-restricted will be converted into a standard, non-proprietary format (ASCII text file) and combined with necessary meta data in the form of a text document and PDF file. Such data will be available on the project website as well as from a professional repository.

General consideration regarding publication, depositing or patenting of research data are summarized by the Figure below that has been reproduced from the H2020 Program Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020:

4   OUTLOOK

LiRichFCC partners are currently reaching increased rates of data generation which make well-crafted policies and processes for data management a must. This report will be distributed among the LiRichFCC partners to focus attention on Data Management issues. At the General Assembly of LiRichFCC at 2017/04/11 in Grenoble, concrete policies and protocols for Data Management will be decided when meeting face-to-face, and the Data Management Plan will subsequently be updated.

5   LIST OF ABBREVIATIONS AND ACRONYMS

ASCII	American Standard Code for Information Interchange
FAIR	Findable, accessible, interoperable, re-usable
GB	Gigabyte = 1,000,000 bytes
LiRichFCC	Project acronym: Lithium-rich face centred cubic [materials]
OpenAIRE	H2020 project and internet platform on Data Management and Sharing
PDF	Portable Document Format