StoRIES

STORAGE RESEARCH INFRASTRUCTURE ECO-SYSTEM

2nd IREP Call Documentation - User

StoRIES IREP1 Post Research Questionnaire (with satisfaction survey): User (s)

After the StoRIES International Research Exchange Programme stay, users are required to submit a post research questionnaire. This should be done within 4 weeks after the Access is completed unless otherwise agreed. The Post Research Questionnaire form will be given to the User(s) by the Project Management. The report contains sections related to the work performed, the main results and observations that were achieved.

[redacted]

Summary questionnaire for users who have been granted International Research Exchnage Programme (IREP) under the StoRIES project Horizon 2020 IREP scheme.

General information about the project	
Project title (as used in application)	Pilot deployment of an **Open** AI/Data Cloud Platform for **Hy**drogen **Mate**rials
Project acronym (max 15 characters)	OpenHyMer
StoRIES IREP RI(s) accessed	*CANMET MATERIALS NRCAN, Hamilton ON, Canada*
Keywords (up to ten, free text)	Data Management, H2 technologies, Ontology, Materials Acceleration Platform (MAP), Self-driving Labs; Grid-scale Storage
Arrival date (in town where RI located)	April 22 2024
Departure date (from town where RI located)	September 1st 2024
Starting date of access (first day at RI)	April 22 to April 30, July 15 to August 31 2024
Finishing date of access (last day at RI)	September 1st 2024
Number of days not using the RI (during the above period)	0

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

Reason for not using RI those days (describe)	
Number of days using the RI	50
Number of users granted access (group size)	50
Comments	
User	
User group leader or sole applicant (user group member 1)	
First name	[redacted]
Last name	[redacted]
Affiliation / Employer	Forschungszentrum Jülich IEK-13
Country of Employer	Germany
E-mail	[redacted]
Comments	
User group member 2	
First name	
Last name	
Affiliation / Employer	
Country of Employer	
E-mail	
Comments	
User group member 3	
First name	
Last name	
Affiliation / Employer	
Country of Employer	
E-mail	
Comments	
User group member 4	
First name	
Last name	
Affiliation / Employer	

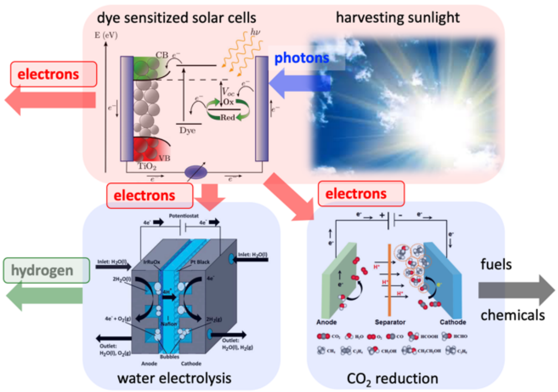
This project has received funding from the European Union's Horizon 2020
Research and Innovation Programme under Grant Agreement N. 101036910

Country of Employer	
E-mail	
Comments	
Please insert more fields if your groups had more than four members.	

Access Summary Report - work performed and initial results

Brief description of the objectives of your project (up to 200 words)

[Please describe short the main objectives of your project]



A critical step in transitioning to a green future is the accelerated discovery and mobilization of high-performance materials for clean technologies. One area requiring rapid advancement to meet global's climate change goals is polymer electrolyte water electrolysis (PEWE) cells. PEWEs are among the most promising water electrolysis technologies, where material innovations are needed due to the high cost of current state of the art (Ir or Ru-based electrocatalysts). The rise of self-driving labs—combining AI with automated robotic platforms—offers a solution, enabling autonomous discovery and development of new materials. Current self-driving labs, however, are limited in scope, often lacking connectivity with conventional labs and broader data management capabilities. There is insufficient integration with external modeling, computational, or experimental data hubs beyond their closed-loop hardware. **The primary objective** of this short visit was to test and pilot utility of an existing cloud platform (ViMi Labs) within several partner labs in Canada (UBC, NRC, NRCan), in particular CanmetMATERIALS. The deployment of few simple models from materials to cells and devices are tested and validated in this pilot. Already being used in few project for data models deployment, this platform aims for automated data management and model deployment to bridge length and time scales (material to device). After completion of this preliminary assessment, the ultimate objective is to support operation of self-driving labs at NRC-EME in Mississauga and NRCan's CanmetMATERIALS in Hamilton, ON.

Activities performed (up to 600 words)

[Please summarise the work carried you (steps taken, instrumentation used, techniques employed, data sources consulted etc.)]

The project plan consisted of the following key activities:

1. **Integration assessment of a data management and model engine platform**: The integration of data management platform ViMi being integrated into a pre-defined workflow. Such workflow consists of a model, input data (extracted from documents or directly obtained from experiment). This assessment aims to streamline data flow between modelling and simulation modules as well as the automated experimental platform.
2. **Models for deployment test:** Identify 1-2 existing materials or cell/device level models for validation, testing and deployment to the platform.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

3. **Benchmarking the performance**: Benchmarking the newly integrated physical and data model against existing workflows. This evaluation will focus on the efficiency, scalability, and accuracy of data handling and processing within automated and connected labs.
4. **Full demonstration of the model deployment and model execution.** developed models will be incorporated into the ViMi Platform, which specializes in R&D data management for energy materials, thus directly enhancing its functionality and utility in ongoing and future research projects. Moreover, the utility of an AI agentic system is evaluated in this task.

Two models have been considered in this project, which are carefully selected, validated and deployed into the platform.

a) Impedance model for performance and degradation predictions of cathode catalyst layer

An impedance model developed was used for modelling the low–Pt cathode catalyst layer (CCL) in a PEM fuel cell as described in detail in [Andrei Kulikovsky 2021 J. Electrochem. Soc. 168 044512, Journal of The Electrochemical Society, 166 (4) F306-F311 (2019)]. The CCL is modeled as a cylindrical pore with a Nafion film separating the open pore volume from the Pt/C surface. This setup aims to characterize the impedance contribution of the Nafion film and the cylindrical pore structure within the catalyst layer. Analytical expressions are derived for the impedance of the Nafion film and the cylindrical pore, as well as for the ohmic resistivity of the CCL, denoted as (R_{ccl}) (in Ohm cm^2). The characteristic frequency of the film impedance is shown to be independent of the film's oxygen transport parameters, complicating the process of distinguishing different transport processes through impedance spectroscopy alone.

A fast algorithm for the distribution of relaxation times calculation is developed, which aids in analyzing and illustrating the impedance related issues, especially in conditions close to the limiting current density in the Nafion film. This is particularly crucial as it manifests the experimental observation of overlinear oxygen transport loss.

b) Predictive CNN-LSTM based model for water electrolysis

The CNN-LSTM model from [Energy and AI 18 (2024) 100420] was selected. It was developed for predicting the performance degradation of a Proton Exchange Membrane (PEM) water electrolyzer incorporates both Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architectures. Here's a detailed summary:

The model is designed for predicting voltage variation over time in PEM electrolyzers, particularly under constant and start-stop operational loads, enhancing the understanding of performance degradation over time. Durability tests were conducted under both constant operation for 1140 hours and start-stop load for 660 hours. The datasets gathered from these experiments were used to train and validate the CNN-LSTM model.

CNN Component: Utilized for extracting relevant features from the input data. The CNN processes the time series data through convolutional layers, identifying important characteristics essential for accurate predictions by the LSTM component.

LSTM Component: This deals with the sequence-to-sequence predictions. It processes the features extracted by the CNN to predict future voltage changes in the electrolyzer, considering the temporal dependencies and patterns observed in the historical data.

The model was trained using the collected datasets, with hyperparameters optimized for best performance. Accuracy evaluations showed R-squared values higher than 0.98, indicating high prediction accuracy compared to other models like LSTM alone and gated recurrent unit (GRU). The model showed its robustness by fitting well to experimental data, and it could make long-term predictions up to 1000 hours with good accuracy.

Scientific results (up to 800 words)

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

[Summarise the (initial) outcomes of your study at the RI(s).]

These activities aim not only to share the advancements made but also to establish new standards in data management for MAPs, fostering greater collaboration and innovation in materials science research. This Pilot deployment of a data management and model engine presents a new and innovative approach to data ontology for accelerated material discovery and process optimization, offering significant advantages over the currently prevalent tabular-database type methods in data management. The parties will benefit from Accelerated materials discovery, especially, requires modular, multi-scale solutions for data management, which is the main goal of this project.

A longer-term scientific impact of this preliminary assessment project is in developing workflow automation and orchestration of modeling tool chains from materials to device integration. The comprehensive user interface aims to utilize data management (autonomous ingestion, standardized data formats and ontologies), modeling and computation workflow automation, orchestrated model execution, and results visualisation across the development pipeline.

To ensure computational workflow automation across scales, labs and domains, the future work will focus on development workflows across the multi-scale framework, and other similar electrochemical systems, including batteries and Solar-to-X materials. One needs to ensure workflows in the platform are modular and adaptable to other technologies. This preliminary project demonstrated a unified data pipeline, integrating and orchestrating modelling and simulation modalities from all different scales. In the future collaboration and for adoption to self-driving labs, we will design the pipeline to accommodate data from various data hubs, including partner labs, internal legacy and conventional as well as relevant self-driving labs regardless of their geo-location.

Interpretation of the results (up to 400 words)

[Discuss the data obtained and describe the major scientific conclusions drawn.]

This pilot project aspires, in long run, to achieve unprecedented levels of scale-to-scale, component-to-component and lab-to-lab integration. It allows a better interpretation of experimental data through the direct deployment of advanced modelling and AI/ML models, filling critical knowledge gaps throughout entire electrochemical systems. This will result in (i) better interpretation of the *operando* behaviour of active materials, catalytic layers, gas diffusion electrodes, and cells/devices, and (ii) effective guidance on further experiments and design of materials, components and cells triggered by the advanced understanding at all scales. This will allow a faster integration of the product improvements into series production at minimal risk and thereby enable shorter (compared to SOTA) development cycles. Data-based correlative approaches already explored for catalyst layers will be utilised and enhanced. Here we focused on models for predicting performance of membrane electrode assembly in PEMWE and PEMFCs. Many parameters influence the performance and longevity of MEAs, including composition and structural properties of components such as catalyst layer (CL), and cell operating conditions. Design parameters extracted from the database included platinum loading (mg/cm^2), platinum to carbon weight ratio (Pt:C) and ionomer content. All these parameters represent the catalyst layer (CL) composition and structure. Although voltage is neither an operational nor a design parameter, its values are a function of operational and design parameters. To ensure the integrity of our predictive modelling, we have excluded maximum power density and similar derived features from the inputs. The preprocessing of the MEA performance dataset, Figure 1, involved several critical steps to ensure its suitability for ML and deep learning analysis.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

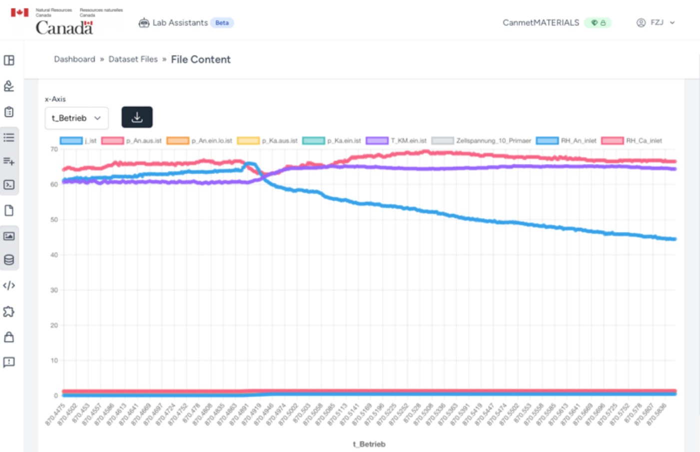


Figure 1. Visual representation of the performance data as input for the predictive MEA model.

The approach in our modelling work has been to create seamless top-to-bottom data workflow structure enabled by a comprehensive framework that will harness (i) extensive synthesis and characterisation data as well as (ii) predictive analytics based on physical-mathematical modelling of relevant processes and materials and accelerated design-build-test-learn (DBTL) cycles for rapid testing and evaluation of new components and cell testing procedures. Figure 2 shows the list of deployed models as being introduced in the previous section.

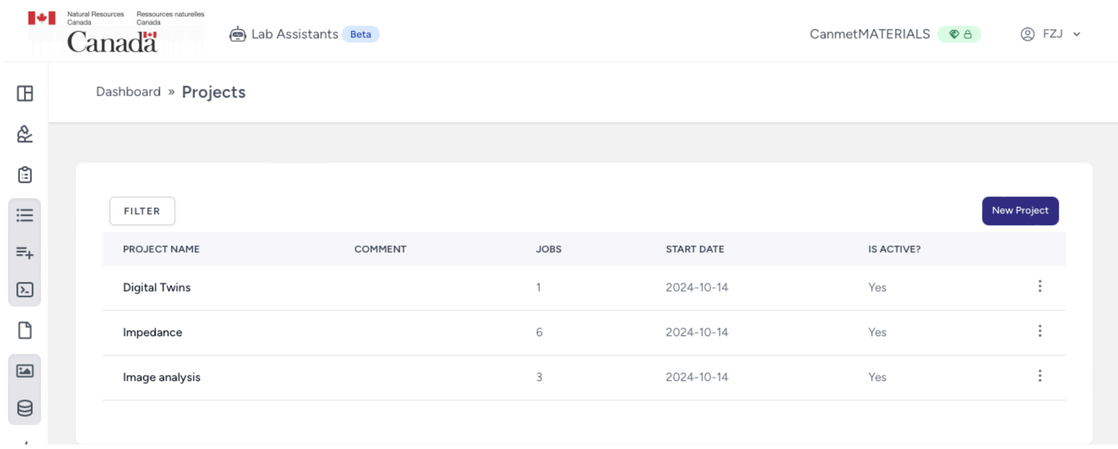


Figure 2. List of cloud deployed models available from the platform for predictive analytics

The impedance model, for instance, provides a simple relation for the Nafion film transport resistivity and compares this to both semi-empirical and model relations available in the literature, enhancing the understanding of oxygen transport mechanisms in low-Pt PEM fuel cells. The models are deployed in the cloud by directly calling the codes from Github repository and adopting to the AWS backend, ensuring model compatibility and execution robustness. Figure

Figure 3 illustrated the input data used for testing the predictive capability of the model, executed directl from the ViMi Platform. The model pinpoints that significant growth in (R {ccl}) occurs near the limiting current density in the Nafion film, around a cell current of about 1 A cm^2, a critical operational parameter for PEM fuel cell designers. This model developed by Kulikovsky highlights the complexity and the interdependencies within the components of the CCL, offering a deeper understanding and several operational predictions that can be quantitatively tested against experimental observations.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

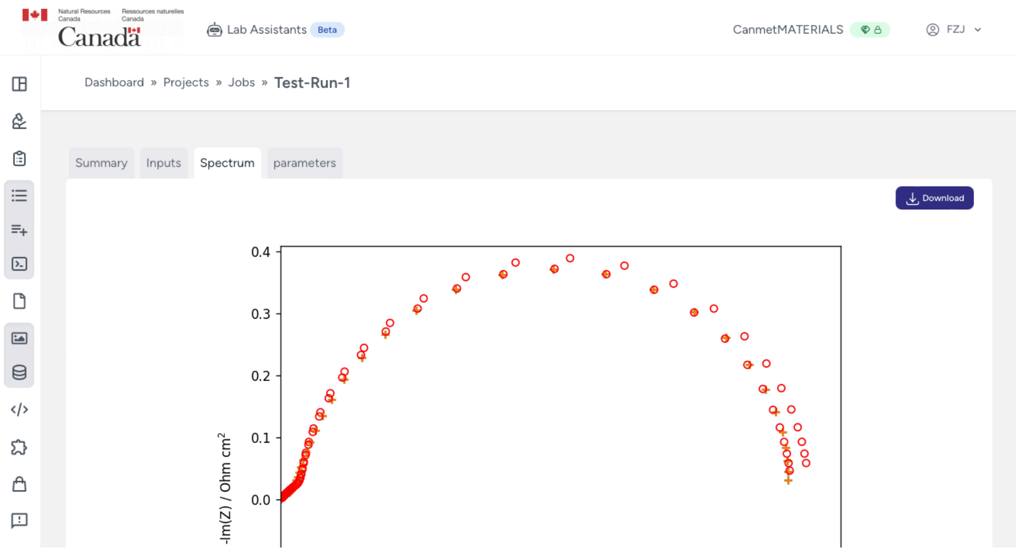


Figure 3. Input data for CL impedance predictive model

The list of extracted data are directly available as part of the analysis dashboard as shown in Figure 4. The resulting parameters are Voltage, CCL resistance, double layer capacitance and oxygen diffusion coefficient.

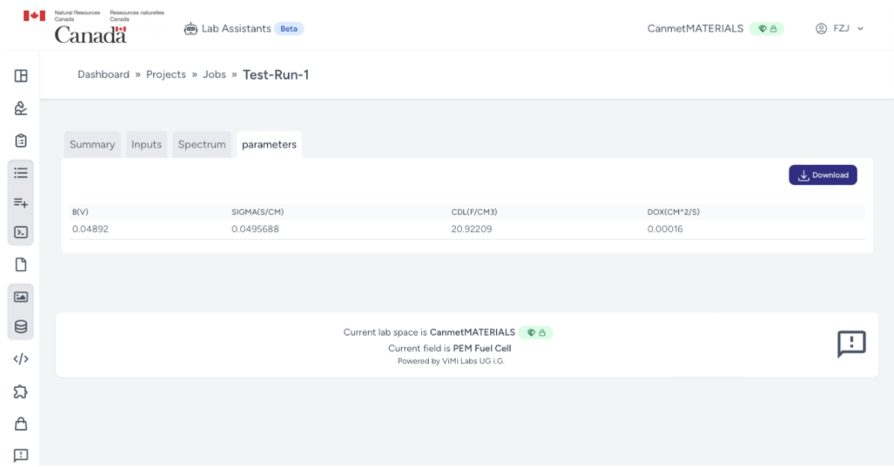


Figure 4. Extracted CCL parameters from the Impedance model

The CNN-LSTM model for PEMWE, on the other hand, demonstrated superior prediction accuracy relative to existing time-series models. As shown in Figure 5, it efficiently handled different operational modes (constant and start-stop), providing reliable degradation rates and predicting future performance based on past operational data. This model serves as a powerful tool for predicting the degradation of PEM water electrolyzers, making it possible to forecast long-term performance and facilitate better operational and maintenance decisions.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

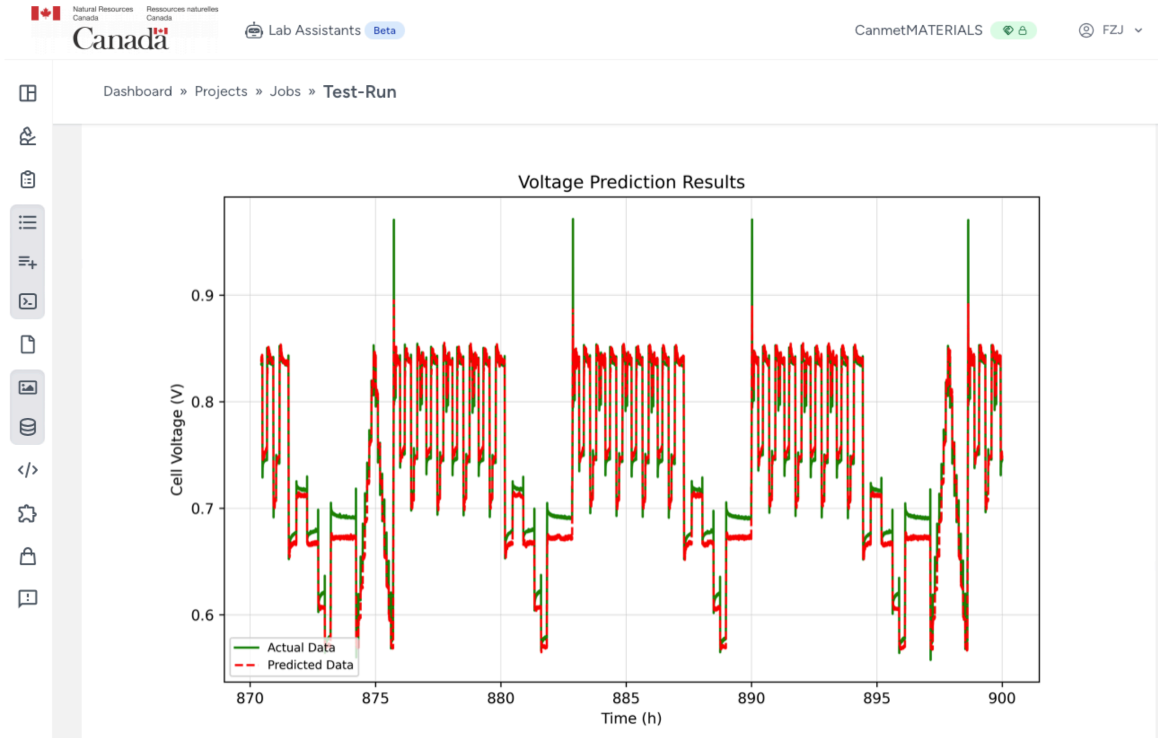


Figure 5. Illustration of predictive power of the CNN-LSTM model for PEMWE, deployed in ViMi platform.

In addition to deploying models and input databases, an agentic system was designed and built, with associated knowledge based. The agents can call databases, call functions and tools and execute accordingly. The performance of few workflows was tested using this agentic system, Figure 6.

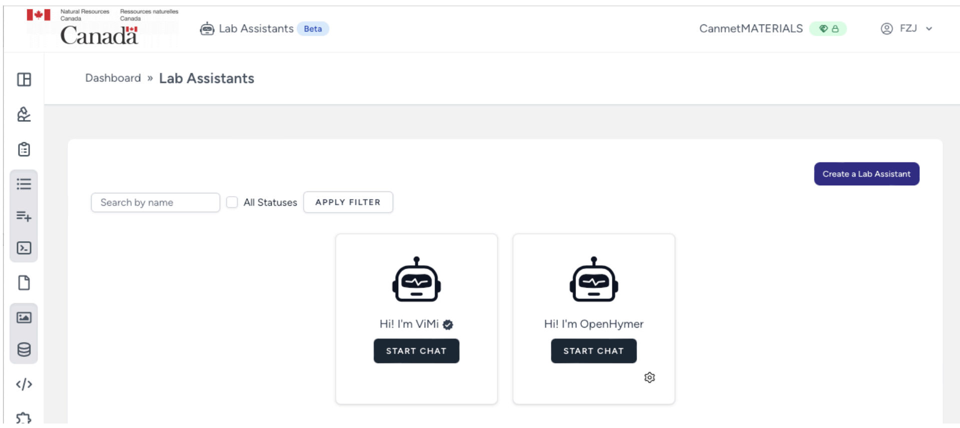


Figure 6. Built-in agentic workflow in ViMi for autonomous orchestration of the modelling pipeline

Main achievements during the IREP related work (up to 250 words)

[Describe the main achievements during your stay at the site(s)]
Successful deployment of simplified ML models in cloud platform, ViMi for prediction of CCL impedance and cell level models. The utility of platform and its agentic workflow was validated for further assessment and adoption t other equipment such as self-driving labs as well as in-operando characterization techniques.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

Difficulties during the IREP related work (up to 250 words)
[List problems and issues you had, completing out your research project: Did you get access to all the necessary equipment, facilities, databases, etc.? If not, please specify the problems that occurred and list equipment the was not working or accessible.] • Delayed with the access to lab space and other infrastructure as a format Lab training and certification was required. • Memory leaks of the predictive models and their cloud compatibility required several trial and testing the server performance • Input databases and extracting the parameter tables from literature for accurate and reliable databases
Intended publications
[Explain where and how you expect to publish the outcomes of your project work. Include also anything already published (What and where?)] The dissemination and exploitation of the project's results will be multifaceted, targeting both academic and industry audiences: 1. Scientific Publication: The findings will be published in a reputable scientific journal, detailing the integration of a graph-based data management platform in Materials Acceleration Platforms (MAPs) and its impact on research efficiency. 2. Conference Presentation: Key results will be presented at the Accelerate Conference, a prominent event in materials science and energy research, to share insights and practical applications with experts in the field.
Data management
[Describe the further usage and storage of project data] The ViMi platform and data management module thereof were used as primary data management system.
Conclusions / additional comments
[Provide any other comments you might have on your work]
Did you complete the EC user questionnaire available at https://ec.europa.eu/eusurvey/runner/RIsurveyUSERS ? (necessary)
☒ *Yes* ☐ *No*

Feedback – HSE, Ethics and Satisfaction

Please rate on a scale from 1 (excellent) to 5 (poor). Feel free to provide additional comments

	1 (excellent)	2	3 (neutral)	4	5 (poor)
Practical information on how to apply for IREP and the overall application process	☒	☐	☐	☐	☐

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

[Comment]	
Information provided, once your project was accepted, on how to proceed	1 (excellent) ☒ / 2 ☐ / 3 (neutral) ☐ / 4 ☐ / 5 (poor) ☐
[Comment]	
Support received at the site(s) regarding technical/scientific matters and logistics	Have you got sufficient support from the RI staff during the project? If not, please, specify the problems. ☒ *Yes* ☐ *No*
[Please specify any problems]	
RI extension / upgrades required	In your opinion, is the RI needed to be upgraded? If yes, please give an explanation. Equipped with API and other cloud connectivity ☒ *Yes* ☐ *No*
[Please specify]	
Problems with local regulations	Have you had any problems with regulations of the visited RI owner (HSE, lab working hours, etc.)? If yes, please, specify ☐ *Yes* ☒ *No*
[Please specify]	
Health and safety issues	Did you encounter any health or safety issue during your research? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	
Environment & Ethics	Did your research involve the use of elements that may cause harm to the environment, to animals or plants? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	
Environment & Ethics	Did your research deal with endangered fauna and/or flora and/or protected areas? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

Environment & Ethics	Did your research involve the use of elements that may cause harm to humans, including research staff? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	
Environment & Ethics – Dual use	Does your research have the potential for military applications? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	
Environment & Ethics – Misuse	Does your research have the potential for malevolent /criminal/terrorist abuse? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	
Environmental issues	Were any potentially dangerous substances (materials / gases etc.) released into the environment (atmosphere, water, or land)? Please provide details. ☐ *Yes* ☒ *No*
[Please provide details]	
Ethics issues	Are there any other ethics issues that should be taken into consideration? Please specify ☐ *Yes* ☒ *No*
[Please provide details]	

Overall impression of communication and interaction after finishing your IREP related work	1 (excellent)	2	3 (neutral)	4	5 (poor)
	☒	☐	☐	☐	☐
[Comment]					

Suggestions for other international and EU facilities not included in StoRIES which you would use for your research
[Please provide suggestions for facilities, describe measurement / experiments you would like to perform]
Suggestions how StoRIES can improve future IREP programme, how to make the IREP more impactful on the energy storage technologies and how to enable the achievement of high TRL levels.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910

[Your suggestions]	
Feedback – Pro-active Innovation Support	
Awareness	Did you know about the pro-active innovation support of StoRIES. ☒ Yes ☐ No
[Please specify how you learned about the pro-active innovation support] *From proposal.*	
Personal experience	Have you taken advantage of or benefited from the pro-active innovation support? ☒ Yes ☐ No
[Please provide details]	
Information/service provided by the pro-active innovation support?	1 (excellent): ☒ / 2: ☐ / 3 (neutral): ☐ / 4: ☐ / 5 (poor): ☐
[Please provide details]	

I declare that the above provided information and especially that information on the number of days visited the RI is correct.

☒ *I have read the StoRIES privacy policy for participation in the StoRIES Accesses and consent to participation and the associated data processing.*

[redacted]

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. 101036910