

DELIVERABLE 4.4
FIRST COURSES/MOOC
AND CURRICULA
DIGITALIZED

BATTERY
2030+

At the heart of a connected green society

Document Classification

Deliverable Number & Title	D4.4.2 First courses/MOOC and curricula digitalized
Work Package	WP4 – European curricula in future battery technologies
Author(s)	DANMARKS TEKNISKE UNIVERSITET (DTU) – Ivano E. Castelli
Delivery due date	2026-08-31
Dissemination Level	Public

Document History

Internal review	2026-08-07	PL Uppsala University Torbjörn Fängström
	2026-08-27	Executive Board
Contributions	23-06-2026	Silvia Bodoardo
	08-07-2026	Simon Clark
Validation	2026-08-31	Project Coordinator
		P1 Uppsala University Patrik Johansson



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957213.

TABLE OF CONTENTS

1	Introduction.....	4
2	MOOC BATTERY COURSE	4
2.1	Rationale and Scope	4
2.2	Course Content	5
2.3	Course Demonstration.....	5
2.4	Hosting on the Battery Genome platform.....	6
3	Digitalisation of curricula	6
4	Conclusions and Outlook.....	7
5	APPENDIX	8

1 INTRODUCTION

Education and training are essential components for accelerating the development of next-generation batteries. The battery field is rapidly evolving, with new materials, cell chemistries, characterisation methods, modelling tools, data standards, and digital workflows being continuously developed. Moreover, most current education programmes cover only a part of the topics needed by battery experts. This creates a strong need for flexible, accessible, and quickly updated teaching material that can be used by students, researchers, engineers, and professionals across Europe.

Digitalised curricula and Massive Open Online Courses (MOOCs) offer a practical way to disseminate common knowledge and best practices across the BATTERY 2030+ community and beyond. They allow training material to be shared beyond individual institutions and projects, ensuring that key concepts, methods, and tools remain accessible also after the end of specific activities. In this way, education becomes part of the long-term infrastructure of the battery community.

In this deliverable, we report the progress towards the development of a digital battery course and digitalised curricula within BATTERY 2030+. Particular emphasis is given to the creation of training material connected to the FAIR and digital battery data activities, represented by the Five-Star Battery Data framework. This framework provides a practical route to publish battery data in a progressively more reusable, interoperable, and machine-actionable form. The corresponding MOOC is designed to translate these principles into practice through examples and reusable resources that can be directly adopted by researchers and students working with battery data. The course is delivered through the Battery Genome Academy and is the first in a planned series of open battery MOOCs that support the wider BATTERY 2030+ effort on FAIR and digital battery data.

2 MOOC BATTERY COURSE

2.1 Rationale and Scope

Battery education requires the integration of many disciplines, ranging from electrochemistry and materials science to data management, modelling, manufacturing, sustainability, and digitalisation. While many courses exist at individual universities, the content is often fragmented and not always available to the broader community including industry. The MOOC developed in this activity addresses this need by providing accessible and structured teaching material on digital battery research. The course is built around the idea that battery data should not only be produced, but also described, stored, shared, and reused in a way that is understandable by both humans and machines. This is particularly important for the current developments in battery activities, where autonomous discovery, digital twins, accelerated testing, and AI-based workflows all rely on high-quality and well-described data. The course

therefore focuses on practical implementation of general principles for data FAIRification. It introduces the motivation for FAIR battery data, the challenges associated with heterogeneous battery datasets, and the stepwise logic of the Five-Star Battery Data framework. The goal is to make researchers able to move from isolated files and local data practices towards structured, annotated, and linked battery datasets.

2.2 Course Content

The MOOC is organised around a set of modules that progressively introduce the main concepts needed to publish and reuse digital battery data.

The first part introduces the role of data in modern battery research and explains why data management is necessary for reproducibility, interoperability, modelling, benchmarking, and long-term reuse.

The second part introduces the Five-Star Battery Data framework. The framework adapts the Five-Star Open Data model to the specific needs of battery research. It guides the user from basic online data publication to structured data, open formats, ontology annotations, and linked data. Each level is presented as a practical step that can be implemented with existing tools and without requiring a complete semantic infrastructure from the beginning.

The third part gives practical examples of how battery datasets can be prepared and published. This includes the description of metadata, the use of persistent identifiers, the choice of open formats, and the connection between data, samples, protocols, contributors, and institutions. The course also introduces relevant vocabularies and ontologies, including schema.org, EMMO, BattINFO, and the Battery Data Format ontology, as tools to make battery data more machine-readable and interoperable.

The final part of the course connects the concepts to real use cases. These examples show how the same principles can be applied to academic research, industrial cell development, product qualification, and large-scale battery testing. In this way, the MOOC does not remain a theoretical description of FAIR data but provides a concrete route for researchers and engineers to apply these ideas in their own work.

The course is openly available through the Battery Genome Academy (<https://www.battery-genome.org/academy>), and the full source material is maintained in a public GitHub repository (<https://github.com/battery-genome/five-star-battery-data>). It is accompanied by a peer-reviewed manuscript, a draft of which is attached as an appendix.

2.3 Course Demonstration

This course has been demonstrated in several activities organised both by BATTERY 2030+ and by individual partners. Around 15 diplomas have been awarded. Feedback has been collected and used to improve the course, in particular the practical exercises. The activities in which the course has been demonstrated and used are listed below:

1. BATTERY 2030+ annual meeting 2025: 100+ participants; target group: battery researchers in academia and industry.
2. SAMBA 2025 (Summer academy for modelling batteries, DTU): 30 participants, target group: PhD students and battery researchers.

3. AI-orchestrated self-driving labs (47216 BSc/47332 MSc courses at DTU): 50 participants, target group: BSc/MSc engineering students.

2.4 Hosting on the Battery Genome Platform

The Five-Star Battery Data MOOC is published and hosted on the Battery Genome platform (<https://www.battery-genome.org>), an open, community-oriented environment for FAIR and digital battery data. Hosting the course on this platform ensures that the training material remains openly accessible, versioned, and directly connected to the data resources, tools, and vocabularies that the course describes.

The Battery Genome platform has been designed as an expandable platform for hosting educational resources. Rather than serving a single course, it provides a common infrastructure into which additional MOOCs and digitalised curricula can be added over time. Each new course can reuse the same publishing, versioning, and dissemination mechanisms, allowing the educational offering of BATTERY 2030+ to grow in a modular and sustainable way as the field evolves.

Building on this infrastructure, a second MOOC dedicated to battery modelling and simulation is currently in development. This forthcoming course will complement the Five-Star Battery Data MOOC by addressing the physics-based and data-driven modelling of batteries, and it will be hosted on the same Battery Genome platform. Together, these courses illustrate how the platform can host a growing family of MOOCs covering the different competences required in modern, digital battery research.

3 DIGITALISATION OF CURRICULA

The MOOC developed within BATTERY 2030+ contributes to ongoing efforts to modernise and digitalise battery education in Europe. By providing accessible, modular and reusable learning materials, the course can be integrated into university curricula, doctoral training programmes and lifelong-learning activities. In this respect, it supports the broader BATTERY 2030+ Education framework (<https://battery2030.eu/education/>), which promotes the development and sharing of educational resources across the battery value chain.

The increasing digitalisation of battery research is creating new educational needs. While electrochemistry, materials science and battery engineering remain fundamental disciplines, modern research increasingly relies on digital tools and data-driven approaches. Researchers are now expected to work with research data management systems, semantic data models, machine-learning methods, modelling environments, digital twins and automated experimentation platforms. As a result, battery education must also address topics such as FAIR data management, interoperability, metadata standards and digital research infrastructures (<https://battery2030.eu/research/research-data-management-rdm-standards/>).

These developments are reflected in several activities carried out within BATTERY 2030+, particularly those related to research data management, battery data standards and the Battery Knowledge Base, as well as semantic frameworks including EMMO, BattINFO and the

Battery Data Format ontology, increasingly adopted within European battery research projects and infrastructures. Together, these initiatives aim to improve the accessibility, interoperability and reuse of battery-related information while supporting the development of digital competences among researchers and students.

The MOOC presented in this deliverable complements these activities by introducing participants to the practical foundations of FAIR battery data. Topics such as metadata management, semantic annotation, persistent identifiers, linked data and battery ontologies are presented in a format that is accessible to learners from different disciplinary backgrounds.

At the industrial level, the European Battery Alliance ecosystem and the InnoEnergy Skills Institute have demonstrated the value of large-scale digital learning platforms for workforce development (<https://innoenergy.com/skillsinstitute/> and <https://www.eba250.com>). Their activities include online training modules, flexible learning pathways, competence-based certification schemes and multilingual educational resources designed to support both upskilling and reskilling within the battery sector. These initiatives illustrate how digital educational formats are becoming an integral component of Europe's battery skills strategy.

Similarly, the Automotive Skills Alliance promotes the development of new educational models for the automotive transition, including battery-related training activities such as the CABATT Battery Course on Automotive Applications, which offers specialised training addressing battery technologies for automotive applications and contributes to the harmonisation of skills development across Europe (<https://automotive-skills-alliance.eu/event/cabatt-battery-course-on-automotive-applications/>).

Additional contributions have been made by projects such as ALBATTs (<https://project-albatts.automotive-skills-alliance.eu/en/home>), which developed competence frameworks and educational models supporting greater harmonisation of battery-related training across Europe. Emerging concepts from BATTERY 2030+ projects, including BATTwin (<https://battery2030.eu/battery2030/projects/battwin/>), further indicate how digital twins, semantic infrastructures and integrated modelling environments may become increasingly relevant components of future battery curricula.

Taken together, these developments point towards the emergence of a common European approach to digital battery education. By combining battery science with FAIR data principles, semantic technologies and digital research practices, the MOOC contributes to preparing students, researchers and professionals for a research environment that is becoming progressively more data-intensive, interconnected and machine-actionable. In doing so, it supports the long-term BATTERY 2030+ vision of a digitally enabled and interoperable battery research and innovation ecosystem.

4 CONCLUSIONS AND OUTLOOK

In this deliverable, we have reported the progress towards the development of a MOOC battery course and digitalised curricula within BATTERY 2030+. The MOOC translates general data-management principles into concrete guidance for battery researchers. It explains how

to publish data using open formats, rich metadata, semantic annotations, persistent identifiers, and linked data concepts. The course therefore supports the broader BATTERY 2030+ objective of making battery research more reproducible, reusable, and machine-actionable.

The digitalised curriculum provides a flexible and reusable foundation that can be integrated into existing university programmes, doctoral schools and professional-training activities, and that can be maintained and extended as battery research evolves.

These activities are part of a broader effort on battery digitalisation. In the future, the Battery Genome platform will host additional MOOCs, including the battery modelling and simulation course currently in development, allowing the digitalised curriculum to expand as new topics and competences emerge.

5 APPENDIX

How to Publish Five-Star Battery Data

Simon Clark,^[a] Lotte Kortstee,^[b] Corsin Battaglia, Christian Punckt, Sergi Obrador, Mukul Parmananda, Benoit Watremetz, Phil Hamilton (Microsoft, TBC), Craig Owens (Microsoft, TBC), Gabe Hege (NVIDIA, TBC), Ivano E. Castelli^[b]

We present a practical framework for publishing battery datasets using linked data standards that enable agentic workflows and adhere to the FAIR guidelines. Drawing inspiration from the World Wide Web Five-Star Open Data model, we outline how battery researchers can publish their data in machine-readable, open formats using shared vocabularies and persistent identifiers. We provide step-by-step guidance on structuring metadata with schema.org, EMMO, and domain-specific battery ontologies, and on publishing using platforms such as Zenodo with semantically rich annotations. Detailed case studies illustrate how to publish time-series data linked to a specific battery cell and test protocol. This approach improves transparency, enables automated discovery, and supports downstream reuse of data from battery research and development environments. This paper is supplemented by a Five-Star Battery Data online course to give engineers, researchers, and students hands-on experience with the concepts presented. The workflow and templates presented here are designed to be immediately adoptable by researchers and engineers working with battery data.

Introduction

Battery research and development is entering a data-rich era. More articles have been published in the last five years than in the previous three decades, while experimental, characterization, and simulation platforms are generating increasingly large and complex datasets. In principle, this rapid growth in data should accelerate progress in battery discovery and engineering. In practice, however, much of the data remains locked in PDF-formatted articles, proprietary binary formats, or local hard drives difficult to access, reuse, or archive^[1].

The main challenge is not only the amount of data being produced, but whether that data can be understood and re-used outside of its original context. Addressing this

challenge requires metadata strategies that go beyond plain-text annotation toward machine-actionable representations of experiments, workflows, and provenance^[2].

Several complimentary approaches for improving digital data publication and exchange have been proposed over the years. The FAIR data principles established a widely promoted framework for ensuring that data are Findable, Accessible, Interoperable, and Reusable. The FAIR principles helped the scientific community evolve from simply making data available upon request toward making it usable by humans and machines by default. Some essential concepts of the FAIR principles include the description of data with persistent identifiers and rich metadata, represented with shared vocabularies with clear provenance and licensing information. These principles have since shaped data management practices across scientific research domains.

In parallel, Tim Berners-Lee and colleagues at the World Wide Web Consortium (W3C) envisioned a future in which machine agents could understand the meaning of data and act autonomously on behalf of people. They called this vision the Semantic Web and developed standards like RDF (Research Description Framework) to transform isolated information silos into a vast network of Linked Data. Although these technologies were accessible to web developers, many of the people generating and publishing data remained unaware of the opportunities for semantic integration and data linking. To bridge this gap, Berners-Lee proposed the Five-Star Open Data model as a simple maturity framework for publishing data on the Web and progressively integrating it into the Semantic Web^[3].

In this model, data publication progresses from making data available online, to using structured formats, to adopting non-proprietary formats, to identifying entities with Uniform Resource Identifiers (URIs), and finally to linking data to other resources. The strength of this model lies in its incremental nature: it does not require researchers or institutions to adopt a complete semantic infrastructure from the outset, but instead provides a clear pathway from basic data sharing toward linked open data. This stepwise logic has been influential because it translates the broader vision of the Semantic Web into practical publication levels that can be understood and implemented by non-specialists.

Throughout the 2010s, the digital infrastructure needed to implement FAIR and Semantic Web principles gradually took shape. General-purpose RDF vocabularies and metadata standards such as Dublin Core, DCAT, schema.org, and CSVW provided reusable building blocks for describing datasets, distributions, tabular data, creators, licenses, access mechanisms, and provenance. These standards made it increasingly possible to publish data with machine-readable context, rather than relying only on human-readable documentation. In parallel, large-scale knowledge graphs such as the Google Knowledge Graph and Wikidata demonstrated the practical value of structured, linked information for search.

[a] Dr S. Clark*
SINTEF Industry, Battery Technology, Strindvegen 4, 7034
Trondheim, Norway
E-mail: simon.clark@sintef.no

[b] L. Kortstee and Prof. I. E. Castelli
Department of Energy Conversion and Storage, Technical
University of Denmark, Agnes Nielsens Vej 301, DK-2800 Kgs
Lyngby, Denmark

[c] M. Parmananda and B. Watremetz
MORROW Batteries ASA