

DELIVERABLE 3.4
FINAL REPORT ON BEST
PRACTICES IN BATTERY
RESEARCH
BATTERY 2030+ CSA3

BATTERY
2030+

At the heart of a connected green society

Document Classification

Deliverable Number & Title	D3.4 Final Report on Best Practices in Battery Research
Work Package	WP3 Common Best Practices in Battery Research Task 3.4: Advocate for the adoption of best practices
Author(s)	P22 EMPA – Corsin Battaglia
Delivery due date	2026-08-31
Dissemination Level	Public

Document History

Internal review	2026-08-07	PL UPPSALA – Torbjörn Fängström
	2026-08-27	Executive Board
Contributions	2026-08-06	P10 SINTEF – Simon Clark
		P5 Fraunhofer ISC – Simon Stier
		P11 DTU – Ivano Castelli
		P7 KIT – Christian Punckt
		Text has been enhanced and refined with the assistance of AI tools to ensure clarity, conciseness, and readability. While the content reflects expert input and has been carefully reviewed, the use of AI has streamlined the editing process.
Validation	2026-08-31	Project Coordinator P1 UPPSALA Patrik Johansson



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant number No. 101104022.

TABLE OF CONTENTS

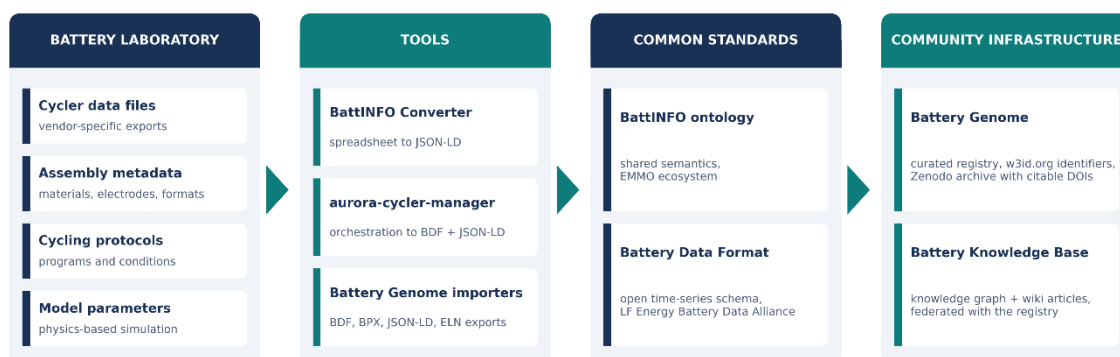
1	Introduction.....	4
2	Report.....	5
2.1	The importance of common standards and best practices in battery research.....	5
2.2	Semantic resources for battery research.....	5
2.3	Lowering adoption barriers with automated annotation tools.....	7
2.4	Open file formats for battery time series data	9
2.5	From data records to community infrastructure	9
2.6	A graded pathway toward FAIR battery data	12
2.7	The ultimate battery dataset	14
2.8	Outreach, training, and community engagement	15
2.9	Sustainability and future work	16
3	Conclusion	17
4	References.....	18

1 INTRODUCTION

To achieve the ambitious goals outlined in the BATTERY 2030+ roadmap,¹ research within the initiative and ideally beyond must adhere to the highest standards for data generation, processing, storage, exchange, and metadata management. A key objective of the initiative is to support the battery research community in developing effective research data management strategies, tools, and consensus-based standards and best practices for both experimental and theoretical research. By integrating research data management tools and standardization, BATTERY 2030+ aims to enhance research quality and to fully implement the FAIR (Findable, Accessible, Interoperable, Reusable) data principles.² This approach has the potential to accelerate battery materials discovery, unlock new battery cell chemistries, foster unprecedented levels of collaboration, and advance data-driven research beyond traditional trial-and-error approaches.

Work Package 3 Common Best Practices in Battery Research of the BATTERY 2030+ CSA3 (grant agreement number 101104022) addresses these needs through four interconnected tasks. Task 3.1 aligns battery domain ontology development and creates training materials on their use. Task 3.2 establishes guidelines for battery data management and sharing. Task 3.3 collects, formalizes, and disseminates standard protocols for theory and experiment. Task 3.4 advocates for the adoption of best practices in the battery research community and supports projects in their implementation through training materials and outreach events.

This deliverable D3.4 is the final report of Work Package 3. It summarizes standards, best practices, tools, and training materials developed across all four tasks and serves as a reference document for the battery community. The report builds on the interim deliverables D3.1 on battery domain ontologies, D3.2 on guidelines for data management and sharing, and D3.3 on standard protocols for theory and experiment, as well as on the results of the preceding CSA2 reported in deliverables D1.5.1 and D1.5.2. Section 2.1 recalls why common standards and best practices matter for battery research. Sections 2.2 and 2.3 describe the semantic resources developed and maintained within the initiative and the tools created to lower their adoption barriers. Section 2.4 presents the open data formats and the collaboration with the international Battery Data Alliance. Sections 2.5 and 2.6 present the community data infrastructure, comprising the Battery Genome and the Battery Knowledge Base, and the Five-Star Battery Data framework, which together translate the FAIR principles into an operational publication pathway. Section 2.7 combines these elements into the description of an ultimate battery dataset and shows how it fulfils the FAIR data principles. Section 2.8 summarizes the outreach, training, and community engagement activities, and Section 2.9 discusses the sustainability of the developed resources beyond the project.



Training and community engagement: online course · documentation · monthly office hours · workshops · memorandum

Figure 1: Overview of the resources developed and promoted in Work Package 3. Laboratory data enters through annotation and conversion tools, is expressed against the common standards for semantics (BattINFO) and time series data (Battery Data Format), and is published through the community infrastructure.

2 REPORT

2.1 The importance of common standards and best practices in battery research

Standards and best practices are crucial in application-oriented transdisciplinary research fields such as battery research. They define clear and uniform methodologies for conducting research and thereby ensure that experiments, analyses, and reporting follow established procedures. This uniformity enables researchers to achieve consistent results across different studies and institutions. Reliable reproducibility is a cornerstone of scientific integrity. It allows others to validate findings by replicating experiments under the same conditions, fosters transparency, and builds trust within the scientific community. Reproducibility is furthermore essential for identifying errors, refining methodologies, and advancing knowledge in a systematic manner.

Common best practices are equally important for interoperability. Modern battery research depends on collaboration across disciplines, institutions, and industries. Standards enable interoperability by defining common formats, structures, and communication protocols for data, ensuring that diverse technologies, software, and datasets can interact seamlessly. Interoperability facilitates the efficient sharing and integration of information, which accelerates progress and minimizes duplication of effort. In addition, standards provide a shared framework for evaluating outcomes, methodologies, and technologies across studies. This comparability is vital for benchmarking and allows researchers, industry, and policymakers to make informed decisions and to assess the performance, quality, and safety of new technologies relative to others. Standards furthermore establish safety frameworks for the handling, processing, storage, transport, and end-of-life management of batteries, while supporting commercialization by enabling compliance with regulatory requirements, such as those related to battery safety, sustainability, and traceability.

A key focus of BATTERY 2030+ Work Package 3 has been the development and promotion of data reporting standards to ensure that battery research data are captured in a consistent, structured, and interoperable manner. Such standards define the essential information required to describe battery materials, components, and cells, experimental procedures, and performance data, enabling reliable comparison, reuse, and integration of datasets across laboratories, research platforms, and organizations. By establishing common reporting practices and aligning them with shared data structures and ontologies, Work Package 3 supports the implementation of FAIR principles and provides the foundation for a more connected and collaborative battery research and development community. These efforts accelerate data-driven discovery, prediction, and optimization of next-generation battery technologies while ensuring that the benefits extend beyond individual projects to the broader battery research ecosystem.

2.2 Semantic resources for battery research

The battery community is generating more data and knowledge than ever before. More battery research papers were published in the five years between 2020 and 2024 than in the entire 33-year period between 1987 and 2019.³ The volume of battery knowledge being published begins to exceed the capacity of humans to process it, and the tools traditionally employed for documenting and sharing knowledge, namely freeform text, narrative-style research papers, and static images, have not kept

pace with advances in information technology. Semantic technologies address this challenge by structuring knowledge and linking data in ways that bridge the gap between human understanding and machine processing. By encoding data in structured and machine-readable formats, they reduce the computational overhead of unstructured data processing while enhancing accuracy and reliability, a capability that becomes even more important as artificial intelligence agents emerge in the field.³

Semantic technologies are built on a small number of core principles.³ Persistent unique identifiers reference resources such as datasets, publications, or equipment and ensure precise identification and retrieval. Controlled vocabularies and ontologies standardize terminology and formalize the relationships between terms to improve consistency across datasets. Machine-readable formats such as JSON-LD embed semantic annotations directly within data files. The resulting technology stack provides a hierarchical framework for organizing structured data. At its foundation, the Resource Description Framework (RDF) expresses information as a network graph of triples, which are three-positional statements expressing basic pieces of information. Vocabularies define standardized terms and ensure consistent data interpretation. Ontologies build upon vocabularies by introducing logic and constraints that support reasoning and enable complex queries. Knowledge graphs extend ontologies and vocabularies to create networks containing instances of data mapped back into the ontology framework, and knowledge bases serve as repositories for structured data, combining a knowledge graph with wiki-style articles and multimedia content useful for both humans and machine agents. Pioneering work in bioinformatics, most prominently the Protein Data Bank and the Gene Ontology, demonstrates the immense value of structuring scientific knowledge in this way. Battery research spans a broader and more heterogeneous landscape covering materials chemistry, electrochemical processes, multiscale manufacturing, diverse testing protocols, simulation, and lifecycle analysis, which motivates the flexible and modular approach adopted by the Elementary Multiperspective Material Ontology (EMMO) ecosystem and its battery domain extensions.

The Battery Interface Ontology (BattINFO) forms the semantic backbone of these efforts.^{4,5} BattINFO is a modular, domain-specific ontology originally created within the EU project BIG-MAP and further developed as a resource within BATTERY 2030+ and the global battery community. As part of the EMMO ecosystem,⁶ BattINFO leverages foundational concepts shared across materials science domains, thereby ensuring semantic consistency and enabling interoperability across disciplines and multiscale modelling workflows. The ontology provides specialized modular vocabularies covering chemical substances, characterization methodologies, and electrochemistry. Its implementation adheres to W3C Semantic Web standards. The ontology is published in Turtle format and maintained in public GitHub repositories to support transparent development, version control, and community contributions, and persistent identifiers managed through the w3id.org service ensure stable referencing over time. Pre-compiled JSON-LD context files simplify the creation and integration of linked data across applications, and a SPARQL endpoint supports semantic querying centered on concepts. The development of BattINFO is led by SINTEF with contributions from partners across the BATTERY 2030+ consortium. Empa contributed modules describing the assembly of different cell formats and cycling protocols, thereby extending the ontology coverage from materials and components to complete experimental workflows. From a functional perspective, BattINFO can furthermore serve as a semantic backbone for operational data management plans, encoding data flows, dependencies, and data handover points in machine-readable form, an approach demonstrated within BIG-MAP and applied to efficient data sharing in larger consortia.⁷

The versatility of this semantic framework is best appreciated through practical examples.³ A typical electrolyte formulation such as 1 M LiPF₆ in equal volumes of ethylene carbonate (EC) and ethyl methyl carbonate (EMC) with 2 weight percent vinylene carbonate (VC) is familiar to experts, yet it lacks the precision necessary for systematic comparison and large-scale analysis, since the component abbreviations are non-canonical and the mixing quantities combine amount concentrations, volume ratios, and weight percentages. Translating such a description into a structured semantic model defines each component within a controlled vocabulary and links its composition, properties, and

functional roles through explicit triples, with unit consistency ensured by alignment with the Quantities, Units, Dimensions and Data Types (QUDT) ontology framework for physical quantities. In the same way, an alphanumeric cell identifier such as IFpR2032, defined by the International Electrotechnical Commission and denoting an LFP/graphite lithium-ion round coin cell with dimensions of approximately 20 mm × 3.2 mm, implies meaning based on assumed pre-existing knowledge. The corresponding knowledge graph makes that meaning explicit and maintains traceability from specific batches of active material to specific cells, and because ontology terms link to external resources such as PubChem and Wikidata, machine agents can follow those links to retrieve additional information that is not explicitly defined in the graph. Once battery data and metadata are machine readable, researchers can execute SPARQL queries to retrieve specific test conditions, for example all test cycles in which the voltage drop exceeded a threshold, enabling automated anomaly detection and performance comparison across different cell designs.

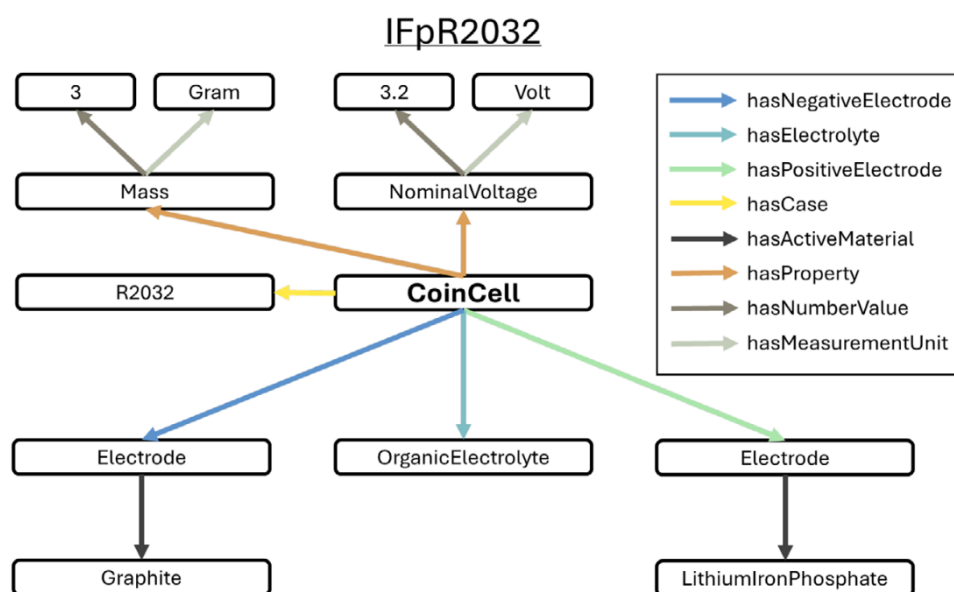


Figure 2: A knowledge graph describing an IFpR2032 coin cell, making the meaning implied by the IEC designation explicit. Reproduced from Ref. 3 (CC BY 4.0).

2.3 Lowering adoption barriers with automated annotation tools

Despite the clear benefits of structuring data in compliance with an ontology, several barriers to widespread adoption exist. Manual semantic annotation of battery data and metadata remains a daunting and error-prone task even for experts, and researchers often lack the technical expertise in ontology-based data management required to implement it in an academic or industrial battery laboratory. To eliminate these barriers, Empa developed BattINFO Converter, an open-source, Python-based web application that automates the conversion of battery cell data and metadata from an Excel table into a structured, semantically annotated JSON-LD format compliant with the BattINFO ontology.^{10,11} The user fills out a familiar Excel metadata template and uploads it to the web application, which performs the conversion without requiring any programming or ontology knowledge and without installing additional software. Because coin cell batteries represent a typical cell format employed by many research labs in Europe, Empa developed a dedicated coin cell battery template. The coin cell template contains ~150 metadata items organized in six categories covering cell identification, positive electrode, negative electrode, electrolyte, separator, and cell assembly, with priorities marked as required, recommended, or optional. Unique online identifiers are used for

[illegible]

In parallel to annotation tools, semantic validation methods play a critical role in ensuring data integrity and reducing human error.³ These include syntactic checks that validate the structure of a record against a schema, semantic consistency checks that flag mismatches such as a length annotated with a unit of mass, and ontology-aware reasoning using inference engines. Frameworks such as SHACL and schema-based tools catch inconsistencies early in the workflow and guide users to correct them. We note that broader adoption still requires balancing machine readability with usability for researchers unfamiliar with ontological frameworks. Ontologies must remain agile enough to incorporate new materials, processes, and testing methods while maintaining internal consistency, which requires active community involvement, an efficient process for creating or modifying terms, continuous updates, and version control to prevent fragmentation. Community-curated examples, detailed documentation, and shared best practices support reproducible use, and user-friendly, artificial intelligence supported interfaces combined with robust semantic validation will be essential for scaling adoption across the battery research community.

CatINFO Converter, an analogous open-source web application for the semantic annotation of CO₂ electrolysis cell metadata based on EMMO's electrochemistry domain,¹³ showing that the semantic framework established by the battery community under BATTERY 2030+ has the potential to propagate also into neighbouring electrochemistry fields.

2.4 Open file formats for battery time series data

Semantic annotation of metadata is most effective when paired with an open, standardized format for the primary data itself. SINTEF and Empa have therefore been working closely with and contributing to the Battery Data Alliance, a Linux Foundation Energy project that develops open standards and tools to accelerate battery innovation and brings together researchers, manufacturers, and developers.²⁰ The Battery Data Alliance develops the Battery Data Format (BDF), an open specification for battery time series data maintained openly on GitHub and released under the Apache 2.0 license.²¹

BDF provides a fixed table schema for battery time series data in which each file contains data for one cell. Three quantities are required in every BDF file, namely test time in seconds, voltage in volts, and current in amperes. Recommended quantities include unix time, cycle count, step count, and ambient temperature. A range of optional quantities covers charging and discharging capacity, step and cumulative capacity, charging and discharging energy, power, internal resistance, applied and ambient pressure, and up to five surface temperature channels. Column headers use preferred labels that follow the IUPAC and SI convention of separating the quantity from the unit with a slash, as in Voltage / V, which preserves algebraic consistency and facilitates automated parsing. Each preferred label is linked to a machine-readable name and to a persistent identifier in the BDF application ontology, which is defined as an extension of BattINFO and thereby ensures interoperability within the broader field of battery data. Files can be serialized to a range of open formats (CSV, JSON, NDJSON, Parquet, Arrow IPC) with optional stream compression (GZIP, XZ, BZ2, ZSTD). The accompanying Python package `batterymdf` provides functions for reading, normalizing, validating, cleaning, plotting, and saving a wide variety of cyclers files into BDF artifacts, both for text-exported and raw binary cycler data. Open formats decouple a dataset's value from the software that produced it, and binding each column to an ontology term decouples the meaning of a dataset from the conventions of a particular laboratory or cycler. Together they allow time series data from different cyclers and laboratories to be parsed, validated, and combined without vendor-specific tooling.

To integrate the open data format with semantic annotation in a single operational workflow, Empa developed a modular open-source software platform for managing and automating massively parallel battery cycling experiments.^{31,33} The `aurora-cycler-manager` software provides a complete pipeline from experiment orchestration to data export. It converts proprietary cycler data from different vendors into BDF, generates BattINFO-compliant JSON-LD metadata for cell assembly and cycling protocols, and packages the results as research-object crates that can be shared through open-access repositories such as Zenodo. The software furthermore supports unified cycling protocol definitions that can be translated into vendor-specific formats for different battery cyclers, thereby eliminating the need to manually rewrite protocols for different equipment. By combining instrument control, data harmonization, and semantic annotation in one platform, this software closes the gap between the annotation tools described in Section 2.3 and the publication infrastructure described in Section 2.5. We consider the combination of an open format governed by an international alliance, a semantic layer governed by the research community, and integrated open-source software tools a robust foundation for scalable FAIR battery data production.

2.5 From data records to community infrastructure

The vision of the Battery Interface Genome within the BATTERY 2030+ roadmap¹ and of the Battery Data Genome initiative¹⁶ both presuppose a shared data substrate in which battery data are uniformly described, linked, and reusable across laboratories. The Battery Genome, developed by SINTEF and

accessible at <https://www.battery-genome.org>, provides an operational, end-to-end implementation of this substrate.¹⁷ The system treats every battery resource as a node in a web of linked data. A small set of record types connected by an explicit provenance chain forms the conceptual core, in which a cell specification describes a class of cells, a cell instance represents a specific physical cell, a test represents an experimental procedure executed on a cell instance, and a dataset captures the data produced by a test. The same specification and instance pattern is applied consistently to materials, electrodes, separators, electrolytes, current collectors, and housings, so that any measured quantity can be traced back to the conditions, protocol, physical cell, and design that produced it. Quantities carry numerical values and explicit units bound to formal ontology terms, and records are validated in layers against JSON schemas, semantic RDF representations, and SHACL shape constraints before publication.

Test protocols receive a two-layer representation. A descriptive layer consists of an ordered list of typed steps such as charge, discharge, hold, and rest, each carrying rates, limits, and termination conditions, and is queryable at corpus scale. An actionable layer links the executable protocol files, for example a cycler program or a simulation experiment definition, which remain the operational source of truth. Users author the descriptive layer using the experiment string syntax popularized by the PyBaMM modeling framework, which the tooling compiles into the structured model and can render back verbatim. Keeping the two layers linked means that a query such as all tests with a C/10 discharge step below 25 °C can be answered over the structured layer, while the exact executable protocol remains attached for reproduction. Knowledge that is reusable across many laboratories, most visibly the specifications of commercially available cells, is captured once in a curated record library maintained through an editorial workflow with an explicit lifecycle, allowing a laboratory testing a widely used commercial cell to reference an authoritative specification and to avoid re-entering the same nominal parameters by hand.

The Battery Genome applies the same living, versioned approach to the best-practice recommendations themselves. A curated guidance collection distinguishes normative practices, which state what a laboratory should do and are reviewed, versioned, citable, and supersedable, from explanatory method guides that describe how to run and analyse individual measurement techniques. Both are scoped to a shared test-kind vocabulary, so that the platform can compose a method page together with the practices that apply to it, and guidance documents reference protocols, records, and datasets in the registry by their persistent identifiers. Because recommendations are maintained under editorial review and revised as techniques and community consensus evolve, the best practices summarized in this report remain living documents on the platform rather than a snapshot frozen at the end of the project.

Data publication separates two roles. The open Zenodo archive acts as the durable store, where every validated data package receives a citable, versioned DOI within the shared battery-knowledge-base community, and a curated registry provides the index over it, minting persistent identifiers under the w3id.org/battinfo namespace and rendering human-readable pages alongside machine-readable Turtle and JSON-LD serializations. Attribution and funding provenance are recorded as machine-actionable metadata on every record. Contributors are identified by ORCID, organizations are reconciled against the Research Organization Registry, funding grants are resolved through the OpenAIRE research graph, and supporting publications are linked through Crossref. Conversion bridges import the formats that laboratories already use, including time series in the Battery Data Format, physics-based model parameter sets in the BPX format, semantically annotated JSON-LD metadata compliant with the BattINFO ontology, electronic lab notebook exports, and executable cycling protocols, thereby allowing incremental adoption. A public discovery platform presents the corpus for human use and is federated with the Battery Knowledge Base, so that every published record links to its knowledge base entry and knowledge base articles reference records by their canonical identifiers. The two infrastructures are complementary, with the Battery Genome holding the curated and

validated record corpus and the Battery Knowledge Base hosting the narrative articles and contextual knowledge that develop around it.

The Battery Knowledge Base, developed by Fraunhofer ISC and available at <https://battery.knowledge-graph.eu>, complements the curated record corpus of the Battery Genome as a community-driven, open-access platform for structured battery knowledge.^{3,8} The Battery Knowledge Base is built on Semantic MediaWiki and the OpenSemanticLab framework⁹ and combines a machine-readable knowledge graph, structured with semantic concepts from BattINFO and EMMO, with traditional wiki-style articles containing human-readable multimedia content. All resources are modelled with ontology-annotated, object-oriented linked data schemas, and identity is managed through ORCID single sign-on so that contributions are associated with persistent researcher identifiers. These schemas follow the Object-Oriented Linked Data (OO-LD) approach (<https://github.com/OO-LD/oold-schema>), in which a single document serves simultaneously as a JSON schema that constrains structure and as a JSON-LD context that supplies semantics, so that one definition drives structural validation, RDF export against BattINFO and general vocabularies such as QUDT, DCAT, and schema.org, and the automatic generation of data-entry forms and programmatic interfaces. The platform hosts entries on processes, products, process parameters, key performance indicators, and measuring techniques along a chemistry-neutral, format-neutral, and scale-neutral battery research and development chain, together with entries on people, organizations, projects, and events that connect the research community. At the time of deliverable D3.3, the Battery Knowledge Base contained over 1400 entries, including more than 200 measuring techniques linked to more than 100 key performance indicators, and it has continued to grow since. Programmatic access is provided through RESTful APIs, a SPARQL endpoint, and a Python package, natural language queries are supported through the Sparklis interface, and the graph is indexed in a retrieval-augmented generation approach that enables direct question answering with large language models.

Mixing Help

Description

mixing of slurry constituent components together with a solvent to create an electrode slurry

Slurry mixing involves combining active materials, conductive additives, binders, and solvents to create a homogeneous electrode slurry mixture^[1]. This slurry is then coated onto a current collector to form the electrode. Proper slurry mixing ensures optimal performance, consistency, and longevity of the batteries.

Components of the Slurry

- Active Material**^[2]: The main component responsible for electrochemical reactions. Examples include lithium cobalt oxide (LCO) and graphite.
- Conductive Additives**^[3]: Materials like carbon black or carbon nanotubes that enhance electrical conductivity.
- Binders**^[4]: Polymers such as polyvinylidene fluoride (PVDF) that provide mechanical stability and adhesion to the current collector.
- Solvents**: Liquids such as N-Methyl-2-pyrrolidone (NMP) or water used to dissolve binders and facilitate mixing.

Mixing Process

A video demonstrating the consistency of lithium-ion battery electrode slurry consistency after different stages of mixing

Mixing

Label: Mixing

Machine compatible name: Mixing

Statements (outgoing):

Statements (incoming):

No results

Category (Class)

Supercategories: [Battery Manufacturing Process](#)

Meta Knowledge Base Category

Battery Manufacturing Process Type

Step ID: 03

Successors: [Casting](#)

Parameters:

- [Batch size \(mixing\)](#)
- [Mixing intensity](#)
- [Mixing order](#)
- [Mixing time](#)
- [Mixing tool \(model\)](#)
- [Mixing tool \(serial number\)](#)
- [Mixing tool \(technology\)](#)
- [Mixing tool \(vendor\)](#)
- [Polymer phase](#)
- [Temperature \(mixing\)](#)
- [Vacuum \(mixing\)](#)
- [Purity \(raw materials\)](#)
- [Particle size distribution \(active materials\)](#)
- [Particle size distribution \(additives\)](#)

Figure 4: An entry in the Battery Knowledge Base describing the mixing of an electrode slurry. The page integrates a human-readable wiki article with a semantic knowledge panel that summarizes the structured metadata and ontological relationships. Reproduced from Ref. 3 (CC BY 4.0).

The Battery Knowledge Base also serves as the vehicle for collecting and formalizing standard protocols under Task 3.3. Formalized process descriptions provide clear guidance for researchers and encourage them to follow and report well-documented, reusable protocols. The BIG-MAP coin cell assembly standard operating procedure and the associated cycling protocol were established as demonstrator projects and are available through the platform, harmonizing experimental practice across

laboratories. As protocols and recommendations gain wider acceptance, the collaborative editing and review features of the Battery Knowledge Base allow the community to collectively improve and update them so that they reflect the latest scientific understanding and align with recommendations from regulatory bodies, funding agencies, and industry groups.

Measured against the operating principles of the Battery Data Genome, the implementation delivers uniform standards and protocols, metadata treated on par with performance data, and systematic curation before downstream use.^{16,17} Module, pack, and system records as well as field data from deployed batteries are not yet modelled, and tiered access with embargo periods for proprietary data remains to be implemented, which will matter for industrial adoption. In relation to the Battery Interface Genome, the infrastructure supplies the uniformly described, linked, and reusable data that the inverse design of battery interphases presupposes, since the contextual information most often lost in publication, such as the formation protocol, the electrolyte composition and batch, and the operating history, is linked structure within the connected data model and can be assembled across laboratories by graph traversal.

2.6 A graded pathway toward FAIR battery data

A gap remains between the general FAIR and Semantic Web principles and the everyday practice of battery researchers and engineers. A researcher who wants to publish a dataset has traditionally had to choose between minimal file deposition and the adoption of a full semantic infrastructure, with little guidance on the steps in between. To close this gap, Work Package 3 developed Five-Star Battery Data, a maturity model that adapts the Five-Star Open Data scheme of Berners-Lee to the practical needs of battery research and translates the FAIR principles into five incremental levels.^{14,15} One star is earned by publishing data online in a repository that issues persistent identifiers, under a clear license such as CC BY 4.0. Two stars require organizing the data in a structured data model with well-defined quantities and units. Three stars require serialization in open, non-proprietary formats such as CSV, Parquet, JSON, or HDF5. Four stars require ontology-based metadata annotations, for example in JSON-LD using BattINFO and general vocabularies such as DCAT and schema.org. Five stars are earned by linking the data to related datasets, entities, and concepts through persistent, resolvable identifiers and by registering the records in a community index. Two deliberate changes distinguish the framework from the original scheme. The fourth star is strengthened from the use of URIs to the inclusion of ontology annotations, because identifiers alone do not convey meaning, and the fifth star emphasizes registration in community indexes, which makes the links between datasets discoverable. Because the levels are cumulative, individual researchers, laboratories, and companies can adopt the framework progressively, starting from whatever practices they have today. The framework is accompanied by a massive open online course that provides hands-on training, worked examples, and reusable templates for each star of the pathway.

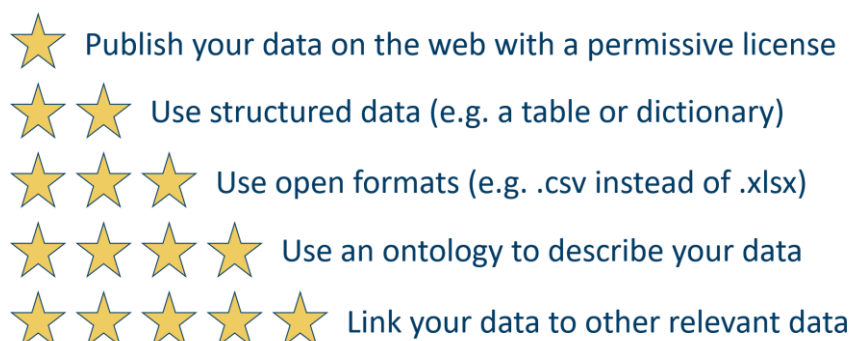


Figure 5: The Five-Star Battery Data pathway, translating the FAIR principles into five cumulative levels from file deposition under a permissive license to fully linked, ontology-annotated data.

For each star, the framework describes what the level requires, how to implement it with existing infrastructure, and what it enables.¹⁴ At the first level, data should be deposited as individual files in repositories that issue digital object identifiers, support versioning, and embed standard bibliographic metadata, under a license that is as permissive as possible and only as restrictive as necessary. At the second level, quantity names and units should adhere to SI and IUPAC conventions to preserve algebraic consistency and to facilitate automated parsing and validation. At the third level, open formats such as CSV for small and medium tables, Parquet for large time series, JSON for nested metadata, and HDF5 for multidimensional arrays ensure that any agent can read the file today and in ten years. At the fourth level, ontology-annotated metadata describes the type of investigation, the object of study, the procedure used, and the outputs produced, with each element drawn from community-agreed vocabularies and serialized in formats such as JSON-LD or RDF Turtle. At the fifth level, the metadata makes explicit references to related resources using well-defined predicates together with machine-resolvable identifiers and human-readable links, and the records are registered in a community registry so that they are indexed, cross-linked, and referenceable by humans and machines alike. Linked, ontology-annotated data furthermore prepares battery datasets for use by artificial intelligence agents, which can only act autonomously if they can discover a dataset, resolve which cell and protocol it describes, verify its license, and retrieve its measurements without human interpretation at each step.

Three case studies demonstrate that this pathway is practical across the battery value chain.¹⁴ The autonomous robotic battery materials research platform Aurora developed by Empa demonstrates an end-to-end implementation of the data pipeline in practice.²³ Aurora integrates electrolyte formulation, coin cell assembly, and battery cell cycling into an automated workflow, with the aurora-cycler-manager software³³ managing experiment orchestration, data parsing, and export in BDF with BattINFO-compliant metadata. A dataset of close to 200 coin cells with NMC/graphite and LFP/graphite chemistries, each tested for 1000 cycles under fully automated and controlled workflows, was published in BDF with fully annotated metadata as the dataset accompanying the peer-reviewed Aurora publication on Zenodo.^{22,23,24} The dataset was co-curated by SINTEF and demonstrates the power of combining robotic assembly, automated workflows, and integrated data management to generate FAIR-compliant battery research data at scale. We consider this release a milestone for Task 3.4, since it demonstrates that data format standardization and semantic annotation developed within BATTERY 2030+ are being adopted and disseminated through an international open-source standardization body represented by the Linux Foundation, reaching well beyond the European battery community.²²

An industrial cell development campaign at Morrow Batteries, comprising 111 released cells of a 100 Ah prismatic LFP/graphite B-sample design tested across temperature, rate, state-of-charge window, and mechanical conditions in 137 test records, shows that industrial data can be published as structured, ontology-ready records that support degradation modelling and benchmarking.

A supplier qualification campaign by Microsoft, comprising 745 LCO/graphite candidate pouch cells for their Surface tablet from three manufacturers in nine model and tab-design groups, demonstrates the complementary perspective of the electronics industry, in which replicate statistics enable supplier comparison as statistical distributions. In this last case study, the complete provenance chain from a published time series back to the cell design it describes can be traversed by machine through persistent identifiers minted under the w3id.org/battinfo namespace. Notably, recomputing the specific energy from the published per-cell mass revealed a single anomalous group whose recorded mass deviated from otherwise identical siblings, which we attribute to a metadata transcription error in the source data. The automated detection of this error by third-party recomputation provides a concrete illustration of the quality control that interoperable, machine-actionable data publication enables.

These industrial datasets illustrate the scientific value that mature data practices unlock.¹⁴ In the Morrow campaign, cells cycled at 10 °C and 25 °C retain roughly 90 to 95 percent of their initial capacity beyond 1400 equivalent full cycles, and replicate cells group tightly within each condition, supporting the use of these fade curves for lifetime model parameterization. In the Microsoft campaign, the within-group scatter of the deliverable energy is typically below one percent of the group mean despite 80 or more replicate cells per group, demonstrating tight cell-to-cell reproducibility and providing a quantitative baseline against which batch outliers or field returns can later be assessed. We note that observations of this statistical quality become shareable and comparable across organizations only when the underlying records are structured, openly formatted, and semantically annotated, which is precisely what the framework promotes.

2.7 The ultimate battery dataset

We conclude the technical part of this report by describing the ultimate battery dataset, an exemplary data package that combines the practices and resources presented in the preceding sections into a single, complete publication. The description serves as a concrete reference against which laboratories can measure their own data publications, identifying which elements are already in place and which remain to be adopted.

The ultimate dataset documents the full experimental provenance of the cells it contains. A first metadata layer describes the cell assembly, covering the active materials, electrodes, electrolyte, separator, current collectors, cell format, and assembly conditions, together with the operator and the institution. This metadata is expressed as ontology-annotated linked data in JSON-LD, with every term drawn from the BattINFO ontology and its EMMO foundation. Chemical substances reference authoritative resources such as PubChem, operators are identified through their ORCID, and institutions through Wikidata or the Research Organization Registry, so that every entity in the description resolves to a persistent, globally unique internationalized resource identifier (IRI). Every quantity carries a numerical value and an explicit measurement unit bound to a formal ontology term.

A second metadata layer describes the cell cycling. The protocol is provided in a human-readable form as an ordered list of typed steps such as charge, discharge, hold, and rest, each with rates, voltage limits, and termination conditions, and in a machine-readable form linked to the executable protocol file that was submitted to the cycler. Environmental conditions such as the test temperature complete the description, so that the experiment can be reproduced exactly and compared systematically with tests performed elsewhere.

The primary data consist of one time series file per cell in the Battery Data Format. Each file contains at minimum the required quantities Test Time / s, Voltage / V, and Current / A, with column headers following the IUPAC and SI slash convention and linked to persistent identifiers in the BDF application ontology. The files are serialized in an open format such as plain-text CSV or binary Parquet, so that any software can read them today and in the future without vendor-specific tooling.

The package is completed by publication metadata. The dataset is deposited in an open repository under a versioned digital object identifier, released under a clear license such as Creative Commons Attribution 4.0 stated in machine-readable form, and registered in a community index where it receives a persistent identifier and links to the cell specification, the protocol, related datasets, the contributing researchers, and the funding grants that supported the work.

Assembled in this way, the dataset fulfils the four FAIR data principles by construction.

Findable. The dataset carries a globally unique and persistent digital object identifier, its rich metadata is indexed in searchable community registries, and every entity within it resolves to its own IRI, so that both the package and its components can be discovered by humans and machines.

Accessible. The data and metadata are retrievable through standard open web protocols from an open repository, without proprietary software or institutional barriers, and the metadata remains available even if access to the primary data should ever be restricted.

Interoperable. The metadata uses the shared vocabularies of BattINFO and EMMO in the standard JSON-LD representation, the time series follow the fixed schema of the Battery Data Format, and qualified references connect the dataset to external resources, so that data from different laboratories can be combined without manual harmonization.

Reusable. The clear license states the conditions of reuse, the complete provenance chain from materials over assembly and protocol to measurement enables informed interpretation, and adherence to community standards ensures that the data meets the expectations of the domain.

Such a dataset is within reach of every battery laboratory today, and a growing family of published datasets realizes the model in practice. The dataset of close to 200 robotically assembled coin cells described in Section 2.6, published in the Battery Data Format with fully annotated BattINFO metadata alongside its peer-reviewed publication, serves as the reference example.^{22,23,24} The approach has furthermore entered regular scientific practice. A recent study on the suppression of fluorine-induced transition metal dissolution in lithium-ion batteries, a standard hypothesis-driven research paper without any methodological focus on data infrastructure, published its complete cycling data in the Battery Data Format with BattINFO-annotated metadata on Zenodo.^{34,35} The flow battery revalidation study described in Section 2.3 published its datasets in the same way.^{12,36} Beyond the consortium, the Discovery Benchmark released by the DIGIBAT facility at Imperial College London provides around 250 lithium-ion coin cells assembled in an automated fashion as a shared benchmark for battery informatics tooling, exported from the datalab platform as a research-object crate together with a version annotated with BattINFO and EMMO ontology terms, demonstrating that the ontology is being adopted internationally even where the Battery Data Format is not yet in use.^{32,37} These examples can serve as guiding examples for future data publications across the community.

2.8 Outreach, training, and community engagement

Recommendations for best practices must be adopted by the research community to have a meaningful impact. Throughout the project, the Work Package 3 team therefore engaged the community through presentations, workshops, webinars, and conferences, soliciting input and feedback and providing training.

The BATTERY 2030+ Memorandum on Research Standards and Guidelines remains an important instrument for community commitment.²⁵ The memorandum was published in July 2021 and can be signed on the BATTERY 2030+ website. Endorsements grew from 91 people from 47 affiliations at the end of the CSA2 in August 2023 to 201 people from 82 affiliations at the end of the CSA3, and the number has continued to grow as new projects join the initiative.

The topics of ontologies, standards, and best practices were presented at BATTERY 2030+ Collaboration Board meetings to introduce the newly starting projects of the initiative to the available resources. The EUnified Battery Data Workshop organized by BIG-MAP in January 2024 brought together over 50 participants from projects across BATTERY 2030+ to discuss a common framework for a European battery data space. At the fourth BATTERY 2030+ Annual Conference in Grenoble in May 2024, with over 300 participants, the work package presented its results in a plenary talk and a poster. The joint BATTERY 2030+ and HYDRA Standards and Best Practices Workshop in Oslo in June 2024 attracted over 80 participants and included a live demonstration of the Battery Knowledge Base. A dedicated Battery Knowledge Base webinar attracted close to 100 registrations, and its public recording had accumulated almost 1500 views. The DigiBatt and IntelLiGent Symposium on Emerging and Digital Battery Technologies, organized by SINTEF, took place in Oslo in June 2025, hosted by the two EU-funded projects.²⁶ The in-person symposium brought together leading battery experts from

research and industry with talks and panel discussions on testing methodologies, data handling, analytics, simulation, and digital twins, complemented by a poster session, and community best practices for the future of battery development were an explicit theme of the program. The DigiBatt Symposium 2026, organized by SINTEF, took place in Oslo in June 2026, again dedicated to digital battery technologies and to the best practices developed within the initiative.²⁷ The Swiss Battery Days 2026, organized by Empa, took place in Zurich in August 2026 and featured a dedicated Battery Data Day.²⁸ The Battery Data Day brings together for the first time all automated battery cell assembly platforms developed within the Horizon projects BIG-MAP²⁹ and BATTERY 2030+, namely the Aurora platform developed by Empa,²³ the ODAcell platform developed by Uppsala University,³⁰ and the robotic platform developed at the University of Muenster, together with the DIGIBAT platform developed by Imperial College London within the Faraday Institution. The program pairs the hardware platforms with the digital infrastructure they feed, including the Battery Genome platform¹⁷ and the Battery Knowledge Base.⁸ We consider this format a template for future community events, since it demonstrates the adoption of the semantic and data infrastructure developed within BATTERY 2030+ as the connective tissue across European automated experimentation platforms.

Training materials accompany all major resources developed in the work package. These include the Five-Star Battery Data online course with worked examples and reusable templates for each level of the framework,¹⁴ the documentation and learning material of the Battery Genome published at battinfo.org,¹⁷ the tutorials and documentation for the Battery Knowledge Base, the documentation of the BattINFO Converter web application with filled example templates, which has accumulated almost 2200 views,¹⁰ and the recorded presentations and webinars available to BATTERY 2030+ members. Together, these materials allow researchers, engineers, and students to adopt the recommended practices without prior experience with semantic technologies.

SINTEF also organizes monthly online BattINFO Community Office Hours, providing a recurring channel for direct engagement with the ontology developers. In these sessions, researchers can raise open issues encountered while adopting BattINFO or its associated tools, ask questions about modeling specific materials, cells, or processes, and receive live guidance on how to extend the ontology for their own use cases. The office hours complement the one-off events described above by offering continuous support throughout the year rather than only around conferences and workshops, and they give the developers a direct channel for gathering feedback that feeds back into the ongoing development of BattINFO. We consider this format an important and lightweight mechanism for lowering the adoption barrier on an ongoing basis, complementing the training materials and tools.

2.9 Sustainability and future work

The resources developed in Work Package 3 are designed to outlive the project. The BattINFO ontology is currently governed by the BATTERY 2030+ consortium and will transition to a community-driven governance model within EMMO after the end of the project. The Battery Data Alliance under the Linux Foundation provides a durable, vendor-neutral home for the continued standardization of the Battery Data Format and related tooling. The Battery Genome, hosted by SINTEF, anchors its published data in the general-purpose Zenodo archive, hosted by CERN, so that published packages remain independently preserved and citable irrespective of the continued operation of any project infrastructure, and its identifiers are operated through the W3C Permanent Identifier Community Group, which decouples them from any single hosting provider. The Battery Knowledge Base is hosted by Fraunhofer. The BattINFO Converter app is hosted by Streamlit. The source code for all projects is available from Github. We note that the long-term funding for data curation and of ontology maintenance remains an open question that the community will need to address together with funding agencies.

A further pillar of sustainability is the alignment of individual incentives with community benefit. Because every record published through the Battery Genome carries machine-actionable attribution

keyed to ORCID and funding provenance resolved through OpenAIRE, researchers accumulate a citable, machine-readable record of their data contributions, and funders can trace the data output of a grant without additional reporting, which responds directly to the incentive problem identified by the Battery Data Genome initiative.¹⁶ In addition, the value of linked data grows with the number of links, so the path forward is as much social as technical. Curating authoritative reference records, integrating with instrument and electronic lab notebook workflows, deepening the federation between the community infrastructures, and promoting the memorandum will remain priorities for the community after the end of the project.

Future work will focus on expanding ontology coverage as new materials, processes, and testing methods emerge, on integrating the knowledge base and the publication pipeline with electronic lab notebooks and instrument-native data formats, on extending the metadata templates of BattINFO Converter to additional cell formats, and on developing intuitive, artificial intelligence supported interfaces that allow researchers and autonomous agents to engage with structured data more easily. As funding agencies and regulatory bodies increasingly emphasize the FAIR principles and as the digital battery passport comes into force, robust semantic frameworks will be needed to maximize the impact and usability of battery data.^{18,19} We furthermore see these community infrastructures as building blocks for the emerging federated data spaces for materials research, such as the European Materials Commons initiative (<https://materialscommons.eu>), in which the same ontology-annotated, object-oriented linked data schemas can carry battery datasets across institutional boundaries and keep them interoperable with the wider materials science and engineering ecosystem.

3 CONCLUSION

In conclusion, Work Package 3 of the BATTERY 2030+ CSA3 has established a coherent set of common standards and best practices for battery research, built around complementary pillars. The BattINFO ontology provides the shared semantic framework for describing battery materials, cells, and experiments in a machine-readable format. The BattINFO Converter lowers the barrier to using this framework by generating semantically annotated metadata directly from familiar spreadsheet templates. The Battery Data Format, now maintained by the Linux Foundation through the Battery Data Alliance, gives the community an open and vendor-neutral standard for reporting cycling data. The Battery Knowledge Base has grown into a reference for the battery community in the spirit of Wikidata and Wikipedia, combining a structured, machine-readable knowledge graph with human-readable articles. The Battery Genome completes the picture as the discovery layer for the community, indexing datasets with persistent identifiers and linking them to the cells, protocols, and people behind them in much the same way a search engine indexes the web, while the underlying data itself is archived on the open repository Zenodo. The Five-Star Battery Data guide ties these pillars together in practice, showing researchers and laboratories how to progress step by step from an isolated file on a hard drive to a fully linked, ontology-annotated, and openly published dataset. Together, these resources translate the FAIR principles into daily practice across the community. A growing number of researchers and organizations, spanning academic laboratories, industrial partners, and international standardization bodies, are adopting these standards, as illustrated by the datasets and publications described in this report. This adoption paves the way for a battery research ecosystem in which data are findable, accessible, interoperable, and reusable by default, thereby accelerating innovation, reducing duplication of effort, and ensuring that public and private research investments deliver the greatest possible benefit to society.

4 REFERENCES

- (1) Amici, J.; Asinari, P.; Ayerbe, E.; Barboux, P.; Bayle-Guillemaud, P.; Behm, R. J.; Berecibar, M.; Berg, E.; Bhowmik, A.; Bodoardo, S.; et al. A Roadmap for Transforming Research to Invent the Batteries of the Future Designed within the European Large Scale Research Initiative BATTERY 2030+. *Adv. Energy Mater.* 2022, 12, 2102785. DOI: 10.1002/aenm.202102785.
- (2) Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* 2016, 3, 160018. DOI: 10.1038/sdata.2016.18.
- (3) Clark, S.; Battaglia, C.; Castelli, I. E.; Flores, E.; Gold, L.; Punckt, C.; Stier, S.; Veit, P. Semantic Resources for Managing Knowledge in Battery Research. *ChemSusChem* 2025, 18, e202500458. DOI: 10.1002/cssc.202500458.
- (4) Clark, S.; Bleken, F. L.; Stier, S.; Flores, E.; Andersen, C. W.; Marcinek, M.; Szczesna-Chrzan, A.; Gaberscek, M.; Palacin, M. R.; Uhrin, M.; Friis, J. Toward a Unified Description of Battery Data. *Adv. Energy Mater.* 2022, 12, 2102702. DOI: 10.1002/aenm.202102702.
- (5) BattINFO Ontology. <https://github.com/BIG-MAP/BattINFO> (accessed July 2026).
- (6) De Baas, A.; Del Nostro, P.; Friis, J.; Ghedini, E.; Goldbeck, G.; Paponetti, I. M.; Pozzi, A.; Sarkar, A.; Yang, L.; Zaccarini, F. A.; Toti, D. Review and Alignment of Domain-Level Ontologies for Materials Science. *IEEE Access* 2023, 11, 120372. DOI: 10.1109/ACCESS.2023.3327725.
- (7) Castelli, I. E.; Arismendi-Arrieta, D. J.; Bhowmik, A.; Cekic-Laskovic, I.; Clark, S.; Dominko, R.; Flores, E.; Flowers, J.; Frederiksen, K. U.; Friis, J.; et al. Data Management Plans, the Importance of Data Management in the BIG-MAP Project. *Batteries Supercaps* 2021, 4, 1803. DOI: 10.1002/batt.202100117.
- (8) Battery Knowledge Base. <https://battery.knowledge-graph.eu> (accessed July 2026).
- (9) Stier, S. P.; Gold, L.; Raeder, A.; Popp, M. A.; Triol, A.; Scheuplein, L. The OpenSemanticLab Platform. *Zenodo*, 2023. DOI: 10.5281/zenodo.11355585.
- (10) Plainpan, N.; Clark, S.; Battaglia, C. BattINFO Converter: An Automated Tool for Semantic Annotation of Battery Cell Metadata. *Batteries Supercaps* 2025, 8, e202500151. DOI: 10.1002/batt.202500151.
- (11) BattINFO Converter. <https://github.com/EmpaEconversion/BattInfoConverter> (accessed July 2026).
- (12) Liu, Q.; Kimbell, G.; Ahn, J.; Camenisch, L.; Celik, R.; Echeverria, T.; Virk, B.; Battaglia, C.; Reber, D. Revalidating KCrPDTA as Flow Battery Electrolyte. Submitted, 2026.
- (13) Plainpan, N.; Kimbell, G.; Beiglboeck, F.; Bissattini Alessi, C.; Bozal-Ginesta, C.; Reber, D.; Senocrate, A.; Battaglia, C. CatINFO Converter: An Automated Tool for Semantic Annotation of CO2 Electrolysis Cell Metadata. Submitted, 2026.
- (14) Clark, S.; Kortstee, L.; Battaglia, C.; Punckt, C.; Obrador, S.; Parmananda, M.; Watremetz, B.; Castelli, I. E.; et al. How to Publish Five-Star Battery Data. Submitted, 2026.
- (15) Berners-Lee, T. Linked Data, Design Issues. <https://www.w3.org/DesignIssues/LinkedData.html> (accessed July 2026).
- (16) Ward, L.; Babinec, S.; Dufek, E. J.; Howey, D. A.; Viswanathan, V.; Aykol, M.; Beck, D. A.; Blaiszik, B.; Chen, B.-R.; Crabtree, G.; Clark, S.; et al. Principles of the Battery Data Genome. *Joule* 2022, 6, 2253. DOI: 10.1016/j.joule.2022.08.008.

- (17) Clark, S.; Flores, E.; Battaglia, C.; Stier, S.; Gold, L.; Punckt, C.; Castelli, I. E.; Vegge, T. Building the Battery Genome. Submitted, 2026.
- (18) Regulation (EU) 2023/1542 of the European Parliament and of the Council Concerning Batteries and Waste Batteries. Off. J. Eur. Union 2023, L 191.
- (19) Battery Pass Consortium. <https://thebatteryypass.eu> (accessed July 2026).
- (20) Battery Data Alliance. <https://batterydataalliance.energy> (accessed July 2026).
- (21) Battery Data Format. <https://github.com/battery-data-alliance/battery-data-format> (accessed July 2026).
- (22) LF Energy Battery Data Alliance. LF Energy's Battery Data Alliance Releases Largest-Ever Open Source Battery Data Format (BDF) Contribution. <https://batterydataalliance.energy/news/largest-dataset-release> (accessed July 2026).
- (23) Svaluto-Ferro, E.; Kimbell, G.; Kim, Y.; Plainpan, N.; Kunz, B.; Scholz, L.; Laeubli, R.; Becker, M.; Reber, D.; Kraus, P.; Kuehnel, R.-S.; Battaglia, C. Toward an Autonomous Robotic Battery Materials Research Platform Powered by Automated Workflow and Ontologized Findable, Accessible, Interoperable, and Reusable Data Management. Batteries Supercaps 2025, e202500155. DOI: 10.1002/batt.202500155.
- (24) Dataset for Publication "Toward an Autonomous Robotic Battery Materials Research Platform Powered by Automated Workflow and Ontologized Findable, Accessible, Interoperable, and Reusable Data Management". Zenodo, 2025. DOI: 10.5281/zenodo.15481956.
- (25) BATTERY 2030+ Memorandum on Research Standards and Guidelines. <https://battery2030.eu/research/research-data-management-rdm-standards> (accessed July 2026).
- (26) DigiBatt / IntelLiGent Symposium on Emerging and Digital Battery Technologies, Oslo, 11 and 12 June 2025. <https://battery2030.eu/news/happenings-events/digibatt-intelligent-symposium-on-emerging-and-digital-battery-technologies/> (accessed July 2026).
- (27) DigiBatt Symposium 2026. <https://digibatt.org> (accessed July 2026).
- (28) Swiss Battery Days 2026. <https://swissbatterydays.empa.ch> (accessed July 2026).
- (29) BIG-MAP, Battery Interface Genome Materials Acceleration Platform. <https://www.big-map.eu> (accessed July 2026).
- (30) Yik, J. T.; Zhang, L.; Sjoelund, J.; Hou, X.; Svensson, P. H.; Edstroem, K.; Berg, E. J. Automated Electrolyte Formulation and Coin Cell Assembly for High-Throughput Lithium-Ion Battery Research. Digital Discovery 2023, 2, 799. DOI: 10.1039/D3DD00058C.
- (31) Aurora Cyclor Manager. <https://github.com/EmpaEconversion/aurora-cyclor-manager> (accessed July 2026).
- (32) Datalab. <https://github.com/datalab-org/datalab> (accessed July 2026).
- (33) Kimbell, G.; Kuehnel, R.-S.; Battaglia, C. A Scalable Open-Source Software Platform for Managing and Automating Massively Parallel Battery Cycling Experiments. Digital Discovery 2026, accepted.
- (34) Kim, Y.; Moysiadou, A.; Egorov, K.; Rentsch, D.; Battaglia, C. A Strategy to Prevent Fluorine-Induced Transition Metal Dissolution in Lithium-Ion Batteries. Adv. Energy Sustainability Res. 2025, 6, 2500194. DOI: 10.1002/aesr.202500194.
- (35) Dataset for Publication "A Strategy to Prevent Fluorine-Induced Transition Metal Dissolution in Lithium-Ion Batteries". Zenodo, 2026. DOI: 10.5281/zenodo.19107066.
- (36) Dataset for Publication "KCrPDTA". Zenodo, 2026. DOI: 10.5281/zenodo.20338409.

(37) Hunter, R.; Hartley, N.; Evans, M.; Garcia Verga, L.; Mildner, F.; Smith, B.; Xiong, S.; Feng, J.; Cooper, S. J.; Walsh, A.; Titirici, M. Discovery Benchmark. Zenodo, 2026. DOI: 10.5281/zenodo.20532539.