Toward a Unified Description of Battery Data

Simon Clark,* Francesca L. Bleken, Simon Stier, Eibar Flores, Casper Welzel Andersen, Marek Marcinek, Anna Szczesna-Chrzan, Miran Gaberscek, M. Rosa Palacin, Martin Uhrin, and Jesper Friis

Battery research initiatives and giga-scale production generate an abundance of diverse data spanning myriad fields of science and engineering. Modern battery development is driven by the confluence of traditional domains of natural science with emerging fields like artificial intelligence and the vast engineering and logistical knowledge needed to sustain the global reach of battery Gigafactories. Despite the unprecedented volume of dedicated research targeting affordable, high-performance, and sustainable battery designs, these endeavours are held back by the lack of common battery data and vocabulary standards, as well as, machine readable tools to support interoperability. An ontology is a data model that represents domain knowledge as a map of concepts and the relations between them. A battery ontology offers an effective means to unify battery-related activities across different fields, accelerate the flow of knowledge in both human- and machine-readable formats, and support the integration of artificial intelligence in battery development. Furthermore, a logically consistent and expansive ontology is essential to support battery digitalization and standardization efforts, such as, the battery passport. This review summarizes the current state of ontology development, the needs for an ontology in the battery field, and current activities to meet this need.

1. Introduction

Batteries are essential to technological progress in the 21st century.[1] Across the industrial landscape, designers and engineers need batteries that are cheaper, safer, and more energy dense.[2] The World Economic Forum projects that the annual battery production revenue will grow to 300 billion dollars per year by 2030.[3] This demand from the market is compounded by the ambitious goals of the Paris Agreement on climate change, which require the transition to renewable energy conversion and storage technologies.[4] The development and large-scale production of sustainable high-performance batteries is one of the most intensely pursued technical research topics in the world today. Bold action is needed to meet this challenge. A coordinated and comprehensive effort spanning research and industry must be undertaken to develop battery technologies that achieve the strict performance standards set by the market in a sustainable and cost-effective way.

Battery data plays an essential role in accelerating the development of new materials, cell designs, models, operating protocols, and manufacturing processes.[5] Recent advances in artificial intelligence (AI) methods such as, machine learning (ML) promise to greatly accelerate and improve insights into battery manufacturing, performance, and lifetime.[6–9] AI methods require vast amounts of data to

S. Clark
New Energy Solutions
SINTEF Industry
Sem Sælands Vei 12, Trondheim 7034, Norway
E-mail: simon.clark@sintef.no

F. L. Bleken
Process Technology
SINTEF Industry
Forskningsveien 1, Oslo 0373, Norway

S. Stier
Fraunhofer Institute for Silicate Research ISC
Digital Transformation
Neunerplatz 2, 97082 Würzburg, Germany

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aenm.202102702.

E. Flores, M. Uhrin
Department of Energy Conversion and Storage
Technical University of Denmark
Anker Engelunds Vej 411, Building 301, Kgs. Lyngby DK-2800, Denmark

C. W. Andersen, J. Friis
Materials and Nanotechnology
SINTEF Industry
Høgskoleringen 5, Trondheim 7034, Norway

M. Marcinek, A. Szczesna-Chrzan
Faculty of Chemistry
Warsaw University of Technology
ul. Noakowskiego 3, Warsaw 00-664, Poland

M. Gaberscek
Department for Materials Chemistry
National Institute of Chemistry
Hajdrihova 19, Ljubljana SI-1000, Slovenia

M. R. Palacin
Institut de Ciència de Materials de Barcelona
ICMAB-CSIC
Campus UAB, Bellaterra, Catalonia 08193, Spain

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**
www.advenergymat.de

make meaningful predictions. Due to the sheer volume of data needed, datasets for AI may include data generated by different laboratories or field sources, which may use different formats, terms, or metadata schemas. Although large amounts of relevant battery data exist, they are often distributed across a variety of online repositories, servers, or PCs, making it difficult to find and access.[10] The concept of data interoperability refers to the needs to integrate from different sources and ensure that data can interoperate with applications or workflows for analysis, storage, and processing.[11,12] Implementing reliable approaches for the interoperability of battery data is essential to fully exploit AI workflows in battery research and development.

Far from being a purely academic exercise, the need for battery data interoperability has a real and wide-ranging impact on the battery industry. Industrial trends toward digitalization of batteries and manufacturing processes, such as, digital twins, must efficiently bring together different types of data from many sources into a coherent framework. Furthermore, battery Gigafactories both require and generate an abundance of battery related data. Gigafactories must manage information coming from external suppliers, service providers, and equipment manufacturers. These factories also constantly generate their own wealth of data from sensors in the manufacturing line, cell formation, quality control, and their own research and development laboratories. Achieving data interoperability among these sources is necessary to both manage the complex logistics within the factory and take full advantage of the generated data to optimize the production process and cell design.

As Gigafactory construction ramps up around the world and batteries become one of the most widely researched topics, vast quantities of battery related data are being generated. How can we distil this deluge of data into knowledge and meaningful action? The answer is to create a universal way of describing and sharing battery data, based on a common conceptualization. This conceptualization can be embodied in a machine-readable battery language, containing both terms and relations needed to describe batteries and their data. The language should be fully open-source and free-to-use so that it is accessible to the entire community. The development should be transparent and engage experts, students, and technicians in the battery space to ensure it reflects the current state of the art and responds to the needs of the community.

Consider the following simple example. If a battery cell manufacturer labels a piece of data with the English word "battery," they probably mean a battery cell. However, an EV manufacturer might use the same term to label data from a battery pack. In some countries, data might be tagged with a synonym like "accumulator," or in a language other than English. Similarly, computers searching for data tagged with the term "battery" might not even refer to an energy storage device at all; it could just as well describe a fortified emplacement for heavy guns or a crime involving unlawful physical contact. Humans can often infer the intended meaning from clues in the context or the data, but machines need explicit and unambiguous descriptors to efficiently interpret the meaning of data.

The FAIR Guiding Principles for scientific data were created as a first step to help address these challenges. FAIR aims to enhance the ability of both humans and machines to (automatically) find and reuse data, and it stipulates that scientific data management should adhere to four core principles: Findability, accessibility, interoperability, and reusability.[11] The FAIR principles are not a formal standard and they do not establish specific technical requirements, rather they define guidelines.[12] Each of the core FAIR principles is supported by a defined set of criteria that should be fulfilled, as shown in Table S1, Supporting Information.

Among the criteria for FAIR data sharing are the requirements that: (meta)data use vocabularies that follow FAIR principles, (meta)data include qualified references to other (meta) data, and (meta)data meet domain-relevant standards. Across the battery community, research groups and journals have started to propose recommendations for data checklists,[13,14] benchmarks,[15] and guidelines[16–21] to improve interoperability. Some authoritative bodies maintain battery terminology glossaries to manage battery and electrochemistry domain vocabularies. The "Electrochemical Dictionary", edited by Bard et al., contains definitions for over 3000 terms in electrochemistry and related fields.[22] These initiatives are helpful, but on their own they are insufficient to fully realize the FAIR data vision. The development of community-driven ontologies offer a solution well-suited to address these needs.[23,24]

Put simply, an ontology is a data model that represents knowledge as a map of concepts within a domain and the relationships between the concepts. Often taking the form an annotated machine-readable knowledge graph, an ontology ascribes meaning to data, provides links to other pieces of data, and allows for machine reasoning according to that meaning. An overview of the aspects and applications of an ontology is shown in **Figure 1**. At its core, an ontology is a model of knowledge that is formalized in machine-readable code. Aspects of an ontology include annotations, which provide context and elucidations of concept. Relations are then used to link concepts together into a logical hierarchy. An ontology provides a common conceptualization and consistent vocabulary to describe a domain. That has applications that include (but are not limited to) software design, data interoperability, and automation.

Ontologies have been applied with great success in the life sciences. The most prominent example is the Gene Ontology (GO).[25] The GO is an ontology of biology that provides structured, computable knowledge regarding the functions of genes and gene products.[26] Over more than 20 years, the GO has grown to become the world's largest source of information on the function of genes and a cornerstone of bioinformatics.[27] Although the life sciences offer some of the most well-known examples of applied ontologies, they have been widely developed and applied throughout many fields of science and engineering. One challenge is that, even though there are substantial on-going efforts to develop new ontologies, the results are often piecemeal and isolated. For example, ontologies may be developed for some specific factory, lab, or niche application without engaging with the whole community or making the ontology openly available. As the battery community races to accelerate the discovery of new energy storage materials and chemistries, it can learn from the community-driven development and adoption of free-to-use and open-source ontologies in bioinformatics. There is a clear need for an analogous cheminformatics initiative to develop a comprehensive standard model of electrochemical systems.

Until only recently, the ontologies needed to fulfil the FAIR criteria and enable the large-scale interoperability of battery data
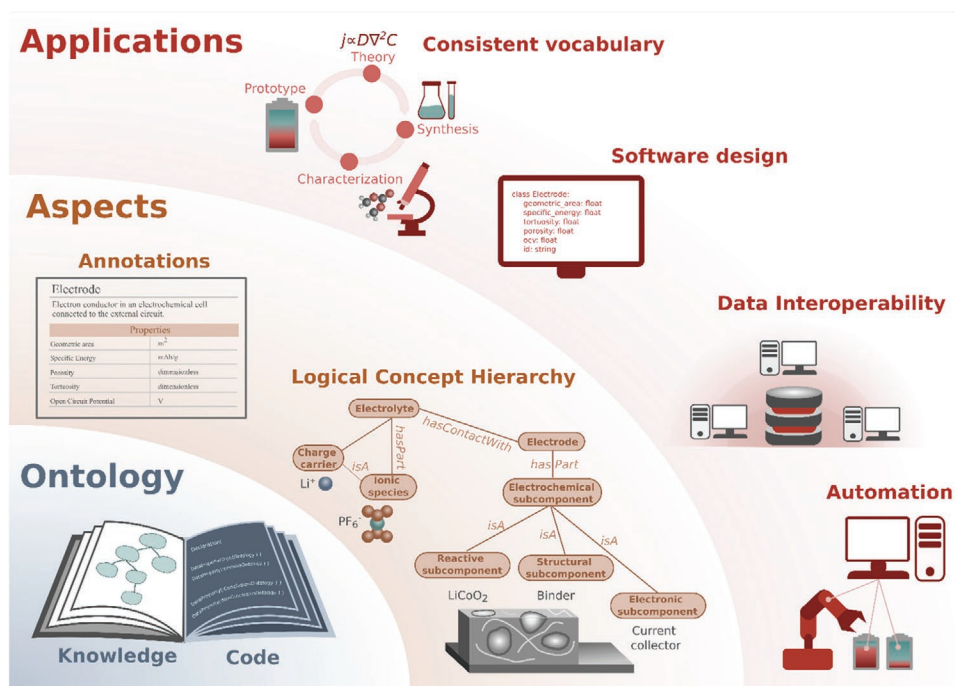
Figure 1. Conceptual overview of an ontology, its aspects, and applications. The core of an ontology is knowledge, which can be expressed as a map of concepts and relations, expressed in machine-readable code. Aspects of an ontology include annotations, which provide context and elucidations of concepts and relations that support the arrangement of concepts in a logical hierarchy. An ontology defines a consistent vocabulary and common conceptualization, which has applications that include software design, data interoperability, and automation.

simply did not exist. Furthermore, many scientists and engineers in the battery community are unfamiliar with ontologies and unaware of the benefits they could bring to their research or product development. This review introduces readers from the battery community to the role of ontologies in the digitalization of battery research and manufacturing. It summarizes the overall role of FAIR data and ontologies in the Semantic Web, reviews the current state of ontology development, and discusses the suitability of existing ontologies to support the growing field of battery research and development.

The remainder of this review is structured as follows. In Section 2, the review begins with a summary of the Semantic Web and introduces W3C concepts related to data interoperability. Section 3 provides an overview of the ontology field by introducing fundamental ontological concepts, an example of a widely used top-level ontology, and a review of current efforts to develop relevant domain ontologies. Section 4 reviews the status of dedicated battery domain ontologies, focusing on two on-going initiatives. Section 5 elaborates on the applications of ontologies in battery research and development, and Section 6 discusses the challenges and outlook for future development. This review may contain terms and phrases that are unfamiliar to some readers new to the field of ontology. The Supporting Information contains a glossary of terms in Table S1, Supporting Information.

2. The Semantic Web of Data

The World Wide Web project was originally developed as a "Web of Documents" to help scientists at the European Organization for Nuclear Research (CERN) share information stored on different devices.[28] The concept used hypertext to link a section of text to some cross-referenced content. Today, the Web has gone on to revolutionize nearly every aspect of how people communicate, and it is populated mostly with content that is designed to be read, understood, and interpreted by humans.

As trends such as, Artificial Intelligence and the Internet of Things continue to grow, there is a need to share large amounts of information that can be interpreted not only by humans but also directly by computers. The World Wide Web Consortium (W3C) is tasked with developing Web standards that address these needs, and it has responded with the concept of the Semantic Web.[29]

The Semantic Web is a "Web of Data" that allows computers to interpret data and do some useful work, much in the same way that a human would interact with the classic Web of Documents. The term semantic refers to the branch of linguistics and logic concerned with meaning. In this context, a piece of data in the Semantic Web is supplemented with metadata that describes its meaning and relation to other pieces of data. By expressing or tagging data with semantic descriptors, computers can interpret the information stored on the web to respond to a query or to achieve a goal.

The Semantic Web is built on a set of fundamental standards adopted by the W3C, including the Resource Description Framework (RDF). The RDF is standard for describing web resources in a flexible and minimally constraining way. A "resource" can refer to anything from a real-world object to an abstract concept if it can be uniquely identified. The RDF uses a three-position statement, known as Triple, to formally express the relationship between two resources.

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
ENERGY
MATERIALS
www.advenergymat.de

An RDF Triple contains three components: A subject, a predicate, and an object. The subject and the object are nodes that represent separable concepts, and the predicate is an arc that describes the relationship between the nodes. The direction of the arc is significant and always point toward the object. For example, a Triple could be a simple statement that a "Li-ion battery" (subject) "is a" (predicate) "battery" (object). Because of the directionality of the "is a" relation, it is known that all Li-ion batteries are batteries but not all batteries are Li-ion batteries.

Each of the components in an RDF Triple are referenced using a uniform resource identifier (URI). A URI is a unique string of characters designed to unambiguously identify a resource. While URIs are limited to US-ASCII characters, the internationalized resource identifier (IRI) is a type of URI that also supports non-Latin characters. Another common type of URI includes the uniform resource locator. The digital object identifier system also allows objects to be represented as a URI.

RDF Triples are essential to create the building blocks of the Semantic Web, known as Linked Data. Linked Data is structured data which adheres to a set of design principles for sharing interlinked data on the web.[30] Linked Data includes both factual data about some specific entities and semantic metadata defining the classes of objects, relationship types, and attributes needed to describe the data. This collection of semantic metadata is collectively referred to as an ontology.

3. Applied Ontology

Ontology is an intricate field that spans philosophy, logic, and computer science. Broadly defined, ontology is the field of philosophy that studies existence, sometimes also referred to as the science of being. Ontologies for computer and information science began coming to prominence in the AI community in the 1970s and 1980s. Whereas normal software development approaches might focus on describing information and concepts, ontologies go deeper to describe things and their natures. AI researchers identified ontologies to both store domain knowledge in a machine-readable format and support automated semantic reasoning about the nature of things using that knowledge. As digital information technologies grew into the 1990s, ontologies were adopted as a layer of knowledge systems to create interoperability standards.[31]

Ontologies play a variety of roles within the scope of the Semantic Web.[24] One role is to assist data integration when ambiguities exist on the terms used in different datasets. For example, when researchers report the fraction of electric current in an electrolyte carried by a certain ion, they might call that quantity the "ion transport number" or the "transference number." Ontologies resolve these issues by assigning a unique identifier, or URI, to a class describing some concept and then annotating that class with information about its definition along with alternative names and links to additional information. Although ontologies can be used simply as technical dictionaries for a domain, their scope extends much further.

Another role of ontologies is to support semantic reasoning to infer new relationships in a dataset, which may yield novel insight to a given problem. Most ontologies applied in computer science are based on so-called description logic, which

has the great advantage of being decidable and ensures that computers can perform logical reasoning within a finite time (given sufficient memory). This allows some ontologies to avoid asserting the full taxonomy in the declaration of the ontology, but rather infer it from logical reasoning. For example, if it is defined that BatteryCell hasPart some ElectrochemicalCell and ElectrochemicalCell hasPart min 2 Electrodes, then it can be inferred that BatteryCell hasPart min 2 Electrodes.

The W3C has developed a dedicated standard called the Web Ontology Language (OWL) to integrate ontologies into the Semantic Web.[32] OWL is a computational logic-based language such that knowledge expressed in OWL can be exploited by computer programs. OWL builds on the RDF concept of Triples but has more extensive and versatile terms for expressing meaning. This allows OWL to express machine-interpretable information that is more semantically rich. An extension of the original OWL language known as OWL 2 was released by the W3C in 2009, and it has since become the most common ontology language in use.[33]

Ontologies can be constructed, shared, and modified using free open-source software. Protégé is an ontology editor and browser developed by Stanford University.[34] It provides a graphical user interface to define ontologies and includes semantic reasoner to make inferences from the asserted relations and axioms. Protégé is available both as a web-based application and as a desktop application and provides a wealth of documentation and tutorials to help orient new users.[35] Owlready2 is a package for ontology-oriented programming in Python.[36] It can load OWL 2.0 ontologies as Python objects, modify them, save them, and perform reasoning via the HermiT[37] and Pellet reasoners.[38] Owlready2 is open-source and available on Bitbucket under a GNU lesser general public license. Extensions of Owlready2 are available such as, EMMO-Python,[39] which includes reasoning with FaCT++[40] and is specifically dedicated to providing tools and functionalities for developing and checking, but not restricted to, EMMO-compliant ontologies.

Through these developments, the definitions of the scope and requirements for ontologies in the technical domain become more formalized. Specifically, the essential points of this definition of ontology are that i) an ontology defines the concepts, relationships, and other distinctions that are relevant for modelling a domain, and ii) the specification takes the form of representational vocabulary (classes, relations, etc.) which provide meanings for the vocabulary and formal constraints on its coherent use.[41]

3.1. Ontological Concepts

An ontology divides a domain into classes that have attributes and are connected by relations. Classes can represent collections of individuals, other classes, or a combination of both. For example, the battery domain ontology might define some classes for *Battery* (encompassing all self-contained contained electrochemical energy storage devices), LithiumIonBattery (describing a subclass of Battery), and GraphiteLCOBattery (a subclass of LithiumIonBattery containing some specific materials). These classes can be related to each other using

Table 1. Primitive elements of ontologies.

Ontological element	Elucidation
Individuals	Distinct basic entities in an ontology. Examples are a specific person, a specific version of a model, a specific dataset, etc.
Classes	The collection of individuals that belong to the class. One can also think about individuals as instantiations of a class, for example, a specific person is an instantiation of the class Person, the specific model is an instance of the class Model, etc.
Relations	Specifications of how classes and individuals are related to each other. An important relation is the isA or isSubclassOf relation, which is used to provide a classification of classes into a hierarchy of subclasses (the taxonomy).
Restrictions	A way to define a class by restricting which individuals that can belong to the class. They are often expressed as a relation combined with an existential, universal, or cardinality requirement.
Annotations	Additional content to the entities in the ontology, without being a part of the logical framework itself. They are very important for making the ontology human understandable.
Axioms	Logical propositions that define the relations between the individuals and classes.

isA relations to create a hierarchy (GraphiteLCOBattery isA LithiumIonBattery isA Battery). This assembly of classes into a logical hierarchy using isA relations is called a taxonomy. Taxonomies are the backbone of ontologies, but ontologies also go beyond simple isA statements to define more relations, restrictions, and axioms that enable a richer description of the system. Some primitive elements of ontologies are described in **Table 1**, and descriptions of some common relations are provided in Table S2, Supporting Information.

One challenge of ontology development is that some classes might need to appear in many different domains. For example, an ontology of the battery domain also needs classes that describe fundamental concepts about physics, chemistry, materials, etc. It is simply not feasible or desirable for a given ontology to encapsulate every single aspect needed to describe some domain. It is more practical to structure ontologies in such a way that they interact with each other and re-use existing class definitions. It is for this reason that one often encounters a hierarchy of ontologies (plural) rather than an ontology

(singular). The development of a hierarchical collection of ontologies that all stem from the same root of basic knowledge and adhere to a set of governing standards can help address this challenge.

Figure 2 illustrates the concept of an ontology hierarchy using the metaphor of a grapevine. This organizational hierarchy is typically divided into four levels: top-level, middle-level, domain-level, and application-level. Top-level ontologies contain the common definitions and theoretical foundations shared by all ontologies in the hierarchy. The top-level ontology is typically small and includes only the essential concepts and well-defined logical theories. For example, definitions of SI units would be included in a top-level ontology. Middle-level (or mid-level) ontologies extend the top-level with concepts that are shared between several domains. For example, the concept of a measurement could be shared by many different domains. In that case, the mid-level ontology might define how a measurement is related to properties, commonly used physical quantities, and units, etc. Domain ontologies describe a particular discipline or
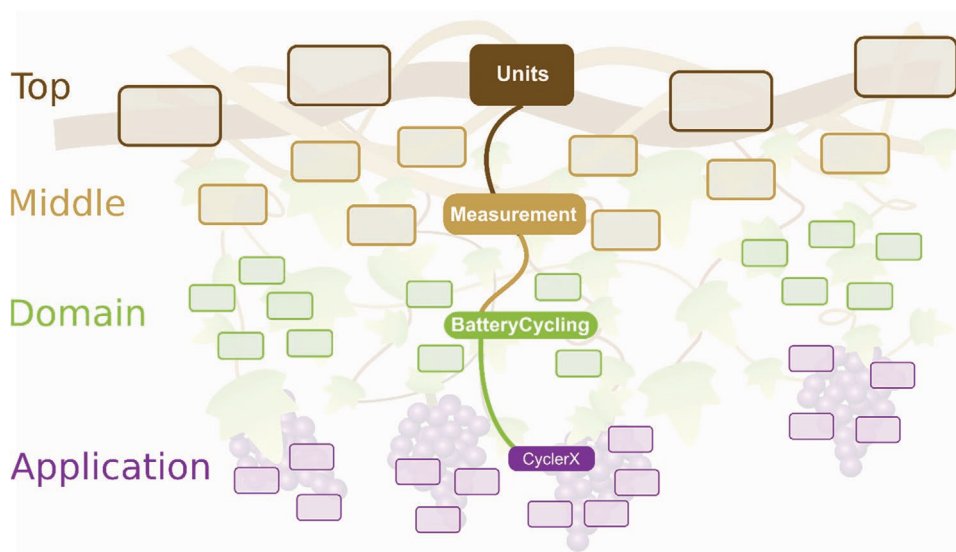

Figure 2. Overview of a hierarchical collection of ontologies, using the metaphor of a grapevine. The top-level contains relatively few core concepts, that are universally applicable, for example, physical units. Middle-level ontologies branch from the top-level and include concepts shared between many domains, for example, a measurement. Domain ontologies extend the middle level with concepts specific to some technical field, for example, a battery cycling measurement. The application ontology is the final level describing some specific use case for the ontology, for example, a specific laboratory with specific measurement equipment.

2102702 (5 of 26)

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**
www.advenergymat.de

field. Domain ontologies provide a common language within the domain, but stay generic, such that they can be used by all relevant applications within the domain. For example, while the mid-level ontology describes the general concept of a measurement, a domain ontology might describe a domain-specific measurement like BatteryCycling. This means that, like top and middle level ontologies, they define classes but rarely individuals. Application ontologies occupy the lowest level and define application-specific concepts and individuals of limited generic interest. For example, the application ontology might contain a description of the specific equipment (CyclerX) used to perform the BatteryCycling measurement. Within all these general categories, there may be sub-hierarchies of ontologies depending on each other.

In 2020, United Kingdom Research and Innovation published A Survey of Top-Level Ontologies[42] where they analyze 37 top-level ontologies to support choosing a suitable top-level ontology for the development of a foundation data model. The ontologies were characterized according to their ontological commitments (e.g., whether space and time are unified or separate, what mereological standard is being adapted, whether they adapt materialism, etc.), formal structure and how they address the individual universal choices. The development and distinctions of top-level ontologies is an extensive topic unto itself, and readers should refer to the above survey for more information. The main activities developing battery domain ontologies today utilize the elementary multiperspective material ontology (EMMO) as the top-level ontology.

3.2. Elementary Multiperspective Material Ontology

The EMMO is a multidisciplinary top- and middle level ontological framework for applied sciences and engineering. EMMO is designed to address the needs for a semantic description which is deeply rooted in the physical sciences, incorporating:[43] i) Description of materials from a rigorous physics perspective; ii) formal relations between granularity levels to facilitate multi-scale materials description; iii) definition of material processes to capture the changing and evolution of materials as chain of different states. The top level of EMMO is very small and concise, only providing the necessary philosophical foundation. A detailed discussion of the structure of the top-level terms in EMMO is provided in the Supporting Information.

The mid-level of EMMO extends the top-level toward domains according to different perspectives. Currently four perspectives are defined: Reductionistic, physicalistic, perceptual, and holistic. The reductionistic perspective introduces a non-transitive direct parthood relation. Whereas a typical parthood relation is transitive (if x is a part of y and y is a part of z then x is a part of z), direct parthood can describe more nuanced situations. For example, we might say that a musician is part of an orchestra and a left arm is part of the musician, but few people would say that a left arm is a part in an orchestra.[44] It provides a powerful way to declare a range of well-defined granularity levels and how to walk between them. It also allows to define concepts like countability and ordering. The physicalistic perspective categorizes the world according to applied physics. Example of subclasses are atom, molecule,

material, and continuum. The perceptual perspective categorizes real world objects according to how they are perceived by a user as a recognizable pattern in space or time. Examples of subclasses include symbols, sounds, and formal languages like metrology or mathematics. The holistic perspective deals with processes that unfold in time and the role of the different participants in the process. Once such process is called a semiosis. In a semiosis—also called a sign process—an interpreter connects an object to some sign that gives it meaning. One example of a semiosis is a researcher (the interpreter) seeing a cat (the object) in their chemistry lab and yelling, "Cat!" (the sign). EMMO utilizes a special theory called Peirce semiotics,[45] which holds that there are three basic semiotic elements (the triadic model) connected by the interpreter: an object, sign, and interpretant. The interpretant refers to the interpreter's internal representation of the object. In the previous example, this is the thought of the cat in the researcher's mind. Within the ontology, semiotic processes are used to define concepts like properties, models, theories, experiments, and observations.

Figure 3 shows a simplified conceptual overview of the multidisciplinary links defined in EMMO. For the purpose of discussion, these disciplines are divided into three conceptual "worlds:" The physical world, the material characterization world, and the modelling world. In the physical world, materials are described according to both their parthood and their physical properties. The material characterization world ontologizes the general process of making a measurement. It uses the holistic perspective to consider a measurement to be a semiosis process an "Observer" perceives another "Object" through a specific perception mechanism and produces some measurement data that is a measure of a physical quantity. That, in turn, can be assigned as a physical property of the object. Finally, in the modelling world, models are defined as being comprised of physics equations and material relations, which themselves contain physical variables that estimate physical quantities.

The development and use of EMMO is supported by the European Materials Modelling Council (EMMC), which aims to bring materials modelling closer to the demands of industry. One core focus of the EMMC is to facilitate the interoperability of data and models across different technical domains. This is done by establishing common metadata schemas and roadmaps,[43] which can be used as a guide for EMMO domain ontologies.

3.3. Elementary Multiperspective Material Ontology Domain Ontologies and Resources

Domain ontologies are designed as extensions of top- and mid-level ontologies to act like plug-and-play libraries of knowledge about a specific topic. They should be specific enough to resolve the necessary information and relationships to describe the topic, but general enough to remain flexible for different applications within the domain.

Consider the example of battery development. To completely ontologize the process of designing and building a new battery, one would need information about not only batteries and electrochemistry, but also models, characterization tools, data management, manufacturing processes, raw material logistics,
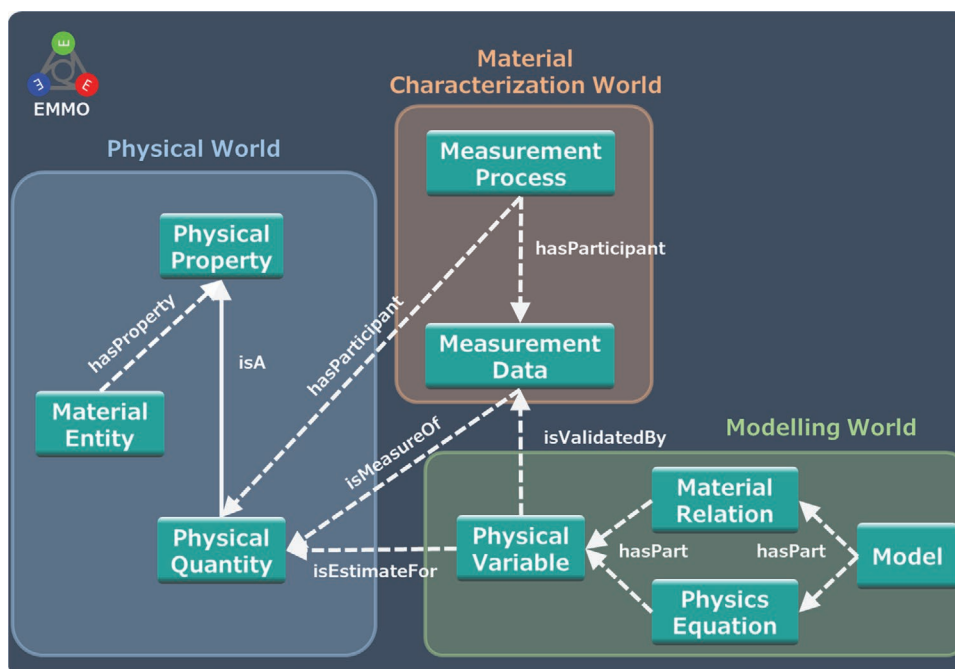
Figure 3. A simplified conceptual overview of the multidisciplinary links between the physical, material characterization, and modelling worlds in defined in EMMO. Terms are shown in blocks connected by arrows indicating relations. Arrows point from the subject to the object.

recycling, and much more. Even with well-developed top- and mid-level ontologies in place, ontologizing every aspect of battery development from scratch would quickly grow to become a herculean task. This challenge can be addressed by a well-managed ecosystem of coherent and consistent domain ontologies. For this reason, EMMC strives after keeping an overview over different initiatives developing EMMO-based domain ontologies (as listed in **Table 2**) and guide their harmonization.

The development of domain ontologies is not an isolated exercise; rather it should be done in close collaboration with relevant domain experts and consider existing standards, checklists, and templates used in the domain community. Before starting a new domain ontology, an ontologist should first check what is already available and determine to what extent existing domain ontologies can be reused. Table 2 presents an overview of public domain and application ontologies utilizing an EMMO top-level ontology. As the EMMO was originally developed to support physics-based modelling activities, a variety of domain ontologies exist to support modelling, simulation, and software development.

In the remainder of this section, we discuss efforts to create EMMO-based domain ontologies for two key topics relevant for battery development: modelling and characterization. More information on other ongoing initiatives developing domain ontologies is available in the Supporting Information.[46–48]

The European Union maintains a Review of Materials Models,[49] which monitors the progress of model development in research and innovation projects funded under the Horizon 2020 program. Now in its sixth edition, the RoMM aims to define and classify the concepts necessary to describe common modelling methods used in industry and research. These definitions form the basis of a harmonized modelling language to help bridge the gaps between theoreticians,

software developers, and end-users. The common modelling language can be used to create metadata definitions that support the goal of generating linked open data in the Semantic Web. The RoMM vocabulary is used as a basis for the Model Data (MODA) template.[50]

MODA is a metadata schema for the standardized description of materials models. It was developed with the goal of helping modelers generate high-level documentation for the models, providing all the information necessary to both reproduce the model and integrate it with other models. Metadata in MODA is defined using plain-text tables and graph notation that is intended to be human-readable. The definition of metadata using the MODA template has become a requirement in many EU-funded research projects to improve model standardization and interoperability. MODA supports modelers in creating standardized high-level documentation that describes all aspects of the model itself and the relevant computational details needed to reproduce the work or interface it to other models.

Although these EMMC initiatives standardizing the metadata have been successful in improving communication and interoperability among humans, they have some shortcomings regarding machine-readable semantic interoperability. Descriptors in MODA are plain-text labels and do not support the integration of linked RDF resources. In MODA, the connections between steps in a modelling workflow are shown using blue arrows, which are ambiguous and can undermine technical reproducibility. The development of an ontological version of MODA could help not only support machine-readable interoperability but also improve the templates by introducing more extensive semantic descriptors. The VIMMP project created an ontology with the goal of addressing this need, called the Ontology for Simulation, Modelling and Optimization (OSMO). The OSMO is an ontological version of the MODA template[51]

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**

www.advenergymat.de

Table 2. Overview of public domain and application ontologies based using the EMMO as top-level ontology.

Ontology	Description
Atomistic and electronic modelling domain ontology	An EMMO-based domain ontology for atomistic and electronic modelling.
Battery interface ontology (BattINFO)	A domain ontology for battery cells and their material interfaces.
Battery value chain ontology (BVCO)	A domain ontology for the battery value chain, including raw materials, manufacturing, and recycling.
Crystallography domain ontology	A crystallography domain ontology based on EMMO and the CIF-ontology. It is implemented as a formal language.
Data model ontology	An ontological description of a simple data model aimed to make application specific data semantic interoperable.
Mechanical testing	A domain ontology for mechanical testing based on EMMO.
General process ontology (GPO)	Describes processes as holistic perspective elements that transform inputs/educts (matter, energy, information) into output/products (matter, energy, information) with the help of tools (devices, algorithms).
Mappings domain ontology	Ontology for mapping to domains and ontological concepts.
Marketplace-accessible computational resource ontology (MACRO)	Ontology for data and hardware related resources and infrastructure.
Materials modelling translation ontology (MMTO)	Ontological formalization of the EMMC translation case template (ETCT).
Mechanical testing domain ontology	A domain ontology for mechanical testing based on EMMO.
Microstructure domain ontology	The microstructure domain ontology is intended to be a domain ontology for physical metallurgy.
Ontology for simulation, modelling, and optimization (OSMO)	Ontological formalization of the EMMC model data (MODA) template.
Ontology for training services (OTRAS)	Ontology including a taxonomy of topics in materials modelling and a formalism by which learning outcomes and expert competencies can be described.
VIMMP communication ontology (VICO)	Ontology describing metadata on messages exchanged at the VIMMP and participants that interact with the platform.
VIMMP software ontology (VISO)	Ontology describing features and properties, including licensing aspects, that can be specified for software packages offered on VIMMP.
VIMMP validation ontology (VIVO)	Ontology addressing model, solver, and processor errors, assessments of computational resource requirements, and benchmarking.
VIMMP ontology of variables (VOV)	Ontology relating physical properties to models, solvers, and the variables occurring in them.

and serves as a metadata standard for modelling and simulation workflows in computational engineering. OSMO is developed with the goal of both making MODA metadata machine-readable and addressing its semantic shortcomings.

Experimental characterization of materials is also the focus of ongoing efforts to standardize data and metadata formats. The Characterization Data (CHADA) template is designed to standardize terminology, classification, and metadata for materials characterization methods.[52] It is analogous to the MODA template for modelling workflows and takes a similar structure, consisting of plain-text tables and graph notation to describe materials characterization workflows. As in the case of the MODA templates, the lack of semantic descriptors in the CHADA template limits its direct use in the creation of Linked Data. Early efforts have been made to ontologize aspects of the CHADA and it is currently under evaluation as source material for a dedicated domain ontology of the EMMO.

Domain ontologies like OSMO and others listed in Table 2 offer some examples that can support the development of ontologies for battery digitalization. In the following section, current efforts to ontologize the battery domain are reviewed.

4. Battery Domain Ontologies

There have been relatively few attempts to develop dedicated ontologies for the battery domain, despite the immense popularity of battery development and the need to support interoperable data. The term "battery" appears in some domain ontologies, but these make no attempt to address the whole of the battery domain itself.[53–55] Dedicated battery domain ontologies can build on the existing initiatives to ontologize materials modelling and characterization workflows described in the previous section, but there is at least one under-developed domain that is needed: electrochemistry.

A domain ontology of electrochemistry can build on existing standards that are widely used and agreed upon in the community.[22] The International Union of Pure and Applied Chemistry (IUPAC) maintains recommendations for terminology. The latest IUPAC recommendations for electrochemical methods of analysis were published in 2020,[56] and serve as an extension of the more general Compendium of Chemical Terminology (also known as the Gold Book).[57] The Electrochemical Society (ECS) has created an online dictionary and encyclopaedia of electrochemistry, which contains over 1000 terms and uses hypertext links to couple entries.[58] The International Electrotechnical Commission (IEC) maintains an active international standard governing the vocabulary of electrochemistry,[59] which is also available as an online digital reference available in a variety of European and Asian languages.[60] While the terminology recommendations from each of these sources are very similar, there do exist some deviations that should be elucidated in a formal ontology of the electrochemistry domain.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**

www.advenergymat.de

A first attempt to represent electrochemistry as an independent domain ontology was proposed by Deepak et al. in 2019.[61] Their ontology includes 18 terms describing different electrochemical cells and reactions. This provides a first introduction and proof-of-concept for how one might approach ontologizing the electrochemistry domain. However, the scope of the ontology is limited. A more extensive description stemming from an established top-level ontology is needed to fully develop an electrochemistry domain ontology.

A potential domain ontology for battery technology can also build on a rich existing standardization landscape. Many battery cell sizes and formats have been standardized for over 100 years. The IEC 60 050 standardizes terminology for primary and secondary cells and batteries and is also available as an online digital reference. The IEC 60 086 series standardizes primary batteries with respect to dimensions, nomenclature, terminal configurations, markings, test methods, typical performance, safety, and environmental aspects.[62] While the IEC standards take precedence for establishing worldwide requirements, the American National Standards Institute (ANSI) maintains battery standards for the United States. The Institute of Electrical and Electronics Engineers (IEEE) standards association has published a glossary of stationary battery terminology.[63] An overview of relevant standards for battery technology is available in Table S3, Supporting Information.

A battery ontology was recently proposed by the German battery research cluster ProZell to target knowledge-based life cycle engineering.[64] They adopted a single-ontology approach for the whole system and has BatteryDataObject as central term, which refers to the Battery being handled, the current LifeCyleStage, the origin of the data object, AnalysisMethod used and TargetInfluence. This ontology is designed to capture life-cycle-oriented information, but also contains basic battery descriptions connected to the BatteryDataObject through the Battery which hasComponent Component. Although it is publicly described, the ontology itself is not yet openly available.

There are currently two major ongoing initiatives dedicated to ontologizing the battery domain: The Battery Interface Ontology (BattINFO) and the Battery Value Chain Ontology (BVCO). BattINFO describes batteries on the cell level and below, including not only components, materials, and their interfaces, but also electrochemical processes, models, and characterization data. The objective of BattINFO is to support AI workflows and interoperability of battery data in the research and development community. On the other hand, BVCO describes aspects of the battery value chain with a strong focus on battery manufacturing and recycling. Both BattINFO and BVCO stem from the top-level ontology EMMO and are publicly available under open-source licenses. The remainder of this section provides an overview of the status of BattINFO and BVCO.

4.1. Battery Interface Ontology

The BattINFO is a domain ontology for battery cells, components, materials, and their interfaces.[65] It is developed with the goal of creating a chemistry-neutral formalized description of batteries to support data interoperability and artificial intelligence workflows. It is defined as a domain ontology of the EMMO, which provides an extensive top-level library of classes and quantities. BattINFO is developed by the BIG-MAP consortium[66] with the short-term goal of supporting data interoperability and artificial intelligence workflows within the battery materials discovery and cell design process.

BattINFO is intended to be a resource available to the whole battery community, free of charge. It is hosted in a public git repository on GitHub and licensed under a creative commons CC-BY-4.0 license. Development of BattINFO began in September 2020. The first version was made public in February 2021, and it is scheduled to have its first stable release in February 2022. This review presents a snapshot of the content and features of BattINFO. Development is on-going, and readers are encouraged to interact directly with the BattINFO source code shared in the public git repository with for instance the free ontology editor Protégé or the python package EMMO-python.

Figure 4 shows an example of a term from BattINFO viewed in the Protégé environment. This example shows the annotations for the Battery entity. The top row shows that the preferred label for the entity (skos:prefLabel) is "Battery," with two alternative labels (skos:altLabel) "ElectricBattery" or the German word "Batterie" that could also be used. Although the primary development of BattINFO is in English, there are planned extensions to support labelling in other languages. The annotations provide links to the dbpedia and Wikipedia entries for the term, and provides the elucidation "One or more electrochemical cells fitted with devices necessary for use, for example case, terminals, marking and protective devices" taken from the IEC standard 60050–482 for electrotechnical vocabulary.[60]

At the heart of BattINFO, a domain ontology for electrochemistry encapsulates fundamental concepts common to all electrochemical systems. Electrochemistry is an intricate and nuanced field. To create a robust and consistent domain ontology, care must be taken to adhere closely to well-established and agreed standards. The foundational classes and elucidations included in the electrochemistry ontology are based on the recommendations of the International Union of Pure and Applied Chemistry (IUPAC)[56] and the International Electrotechnical Commission (IEC). Where necessary, these are supplemented with information from preeminent electrochemistry textbooks.[67–69]

Electrochemistry is often framed within the scope of its coupled and foundational electrochemical components: an electrode, an electrolyte, and an electrochemical cell. It is at the electrode-electrolyte interface that many phenomena determining the behavior of the electrochemical cell occur. The electrodes and electrolytes exhibit intrinsic material properties, which determine observable cell-level quantities.[70] A comprehensive and well-grounded ontological description of an electrode, an electrolyte, and an electrochemical cell are essential to an electrochemistry ontology. Ontological descriptions should, at a minimum, include an i) elucidation of the meaning of the term and information about ii) its properties and iii) relationships with other terms.

Table 3 compares definitions of some foundational terms in electrochemistry from four authoritative sources. Even among these widely respected references, there is some significant variation in the definitions.
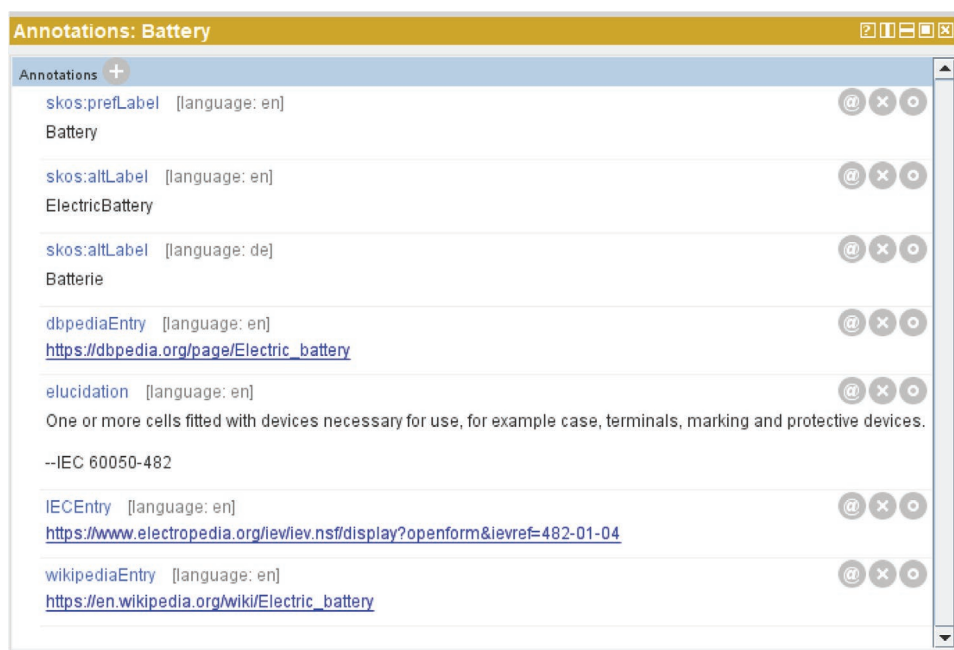
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
ENERGY
MATERIALS
www.advenergymat.de

Figure 4. Example of annotations for an entity in BattINFO as visualized in Protégé.

4.1.1. Electrode

An ontological description of an electrode must first contain a natural language elucidation of the term. Definitions from authoritative sources are compared in Table 3. The IUPAC definition focuses on the ability of an electrode to conduct electrons between the electrochemical cell and the external circuit; a similar definition from Electrochemical Systems highlights that the electrons must be the mobile species in the material. The Handbook of Batteries views the electrode not as a part or material with certain properties but as a physical space at which electrochemical processes take place. Only the definition from the IEC highlights that the electrode must share an interface with a separate medium and exchange charge carriers to function. Recalling that ontological definitions should aim to encompass the nature of things (not only their properties) the definition from the IEC is most appropriate. It is accepted in BattINFO with the clarification that the term "conductive" refers to electronic conductivity. The other definitions are still included as comments in the entity to provide context to users.

Table 3. Definitions of some foundational terms in electrochemistry from authoritative sources.

| Term	Definition	Source
Electrode	An electron conductor in an electrochemical cell connected to the external circuit.	IUPAC[56]
	A conductive part in electric contact with a medium of lower conductivity and intended to perform one or more of the functions of emitting charge carriers to or receiving charge carriers from that medium or to establish an electric field in that medium.	IEC 60 050[60]
	A material in which electrons are the mobile species and therefore can be used to sense the potential of electrons	Electrochemical systems[67]
	The site, area, or location at which electrochemical processes take place.	Handbook of Batteries[69]
Electrolyte	Conducting medium in which the flow of electric current is accompanied by the movement of ions.	IUPAC[56]
	Liquid or solid substance containing mobile ions that render it conductive.	IEC 60 050[60]
	A material in which the mobile species are ions and free movement of electrons is blocked.	Electrochemical systems[67]
	The medium which provides the ion transport mechanism between the positive and negative electrodes of a cell.	Handbook of Batteries[69]
Electrochemical cell	A system that consists of at least two electron conductors (electrodes) in contact with ionic conductors (electrolytes).	IUPAC[56]
	A composite system in which the supplied electric energy mainly produces chemical reactions or, conversely, in which the energy released by chemical reactions is mainly delivered by the system as electric energy.	IEC 60 050[60]
	A system containing two electrodes that allow transport of electrons, separated by an electrolyte that allows movement of ions but blocks movement of electrons.	Electrochemical systems[67]
	The basic electrochemical unit providing a source of electrical energy by direct conversion of chemical energy. The cell consists of an assembly of electrodes, separators, electrolyte, container and terminals. (Definition of "Cells")	Handbook of Batteries[69]
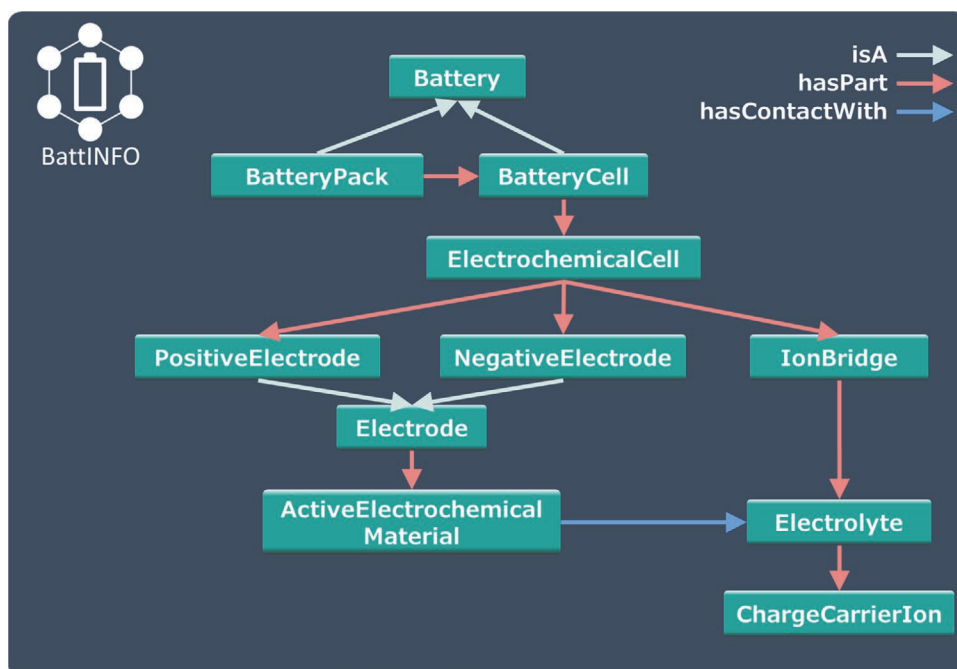
Figure 5. Simplified overview of some core concepts describing a battery in BattINFO. Arrows point from the subject to the object.

Figure 5 shows the basic relations for an electrode in BattINFO. From the elucidation, it is known that the electrode must contain one or more parts that support the conduction of electrons, it must have contact with an adjacent medium (electrolyte), and it must support the exchange of charge carriers across the interface (i.e., via an electrochemical reaction). It can be assumed that the electrode must remain physically intact to function. Therefore, the electrode must contain some part of a material that is electronically conductive and participates in an electrochemical reaction with the adjacent medium. In BattINFO, this material is termed the active electrochemical material.

Ontological descriptions must take care to be sure that the defined concepts and relations are not too restrictive. For example, when asked to describe a battery electrode, many researchers from the Li-ion battery field describe it as a composite having parts current collector, active material, binder, and additives. That is a valid description of one type of electrode. However, a battery electrode could also be a simple piece of zinc foil. BattINFO is designed to be chemistry neutral. Therefore, it is necessary to start by defining general descriptions of the nature of things and adding layers of subclasses to achieve the required level of detail.

Electrodes are often classified according to whether the dominant electrochemical reaction is an oxidation (cathode) or a reduction (anode). For electrically rechargeable electrochemical cells, the electrodes assuming the roles of anode and cathode change between the discharging and charging process. Although the common practice in the battery field is to refer to the electrodes by their function during the discharging process, this introduces a semantic inconsistency in the description. For this reason, BattINFO distinguishes electrodes using the terms positive electrode and negative electrode. The positive electrode is defined as the electrode with the higher electric potential under open-circuit conditions, and it fills the role of cathode during discharging and anode during charging.

4.1.2. Electrolyte

The definition of electrolyte is more consistent across sources, as shown by the comparison in Table 3. Each of the four considered sources highlights that the essential aspect is the presence and movement of ions in the material. The Electrochemical Systems definition includes the caveat that the material must also block the movement of electrons. This definition is accepted in BattINFO. The electrolyte is essential to the function of an electrochemical cell and is therefore classified as an electrochemical component in BattINFO. A simplified overview of some Electrolyte relations in BattINFO is shown in **Figure 6**.

The electrolyte is among the most important and complex materials in an electrochemical system. It can take many different forms including electrolytic solutions (using either aqueous or organic solvents), ion-conducting gels, polymer membranes, or solid-state ceramic materials.[71] However, there are some aspects of electrolyte design that are common to all electrolytes regardless of their exact composition.

Electrolytes are electronically insulating and ionically conducting materials. They must resist the flow of electrons to force them through the external circuit while also shuttling ions to sustain the electrochemical reactions at the electrodes. An electrolyte must contain at least one mobile ion, known as the charge carrier ion, to shuttle charge between electrodes. To maintain charge-neutrality in the electrolyte, the charge carrier ion must be accompanied by a counter ion of opposite charge. The sum of positive charges, either mobile or immobile, must
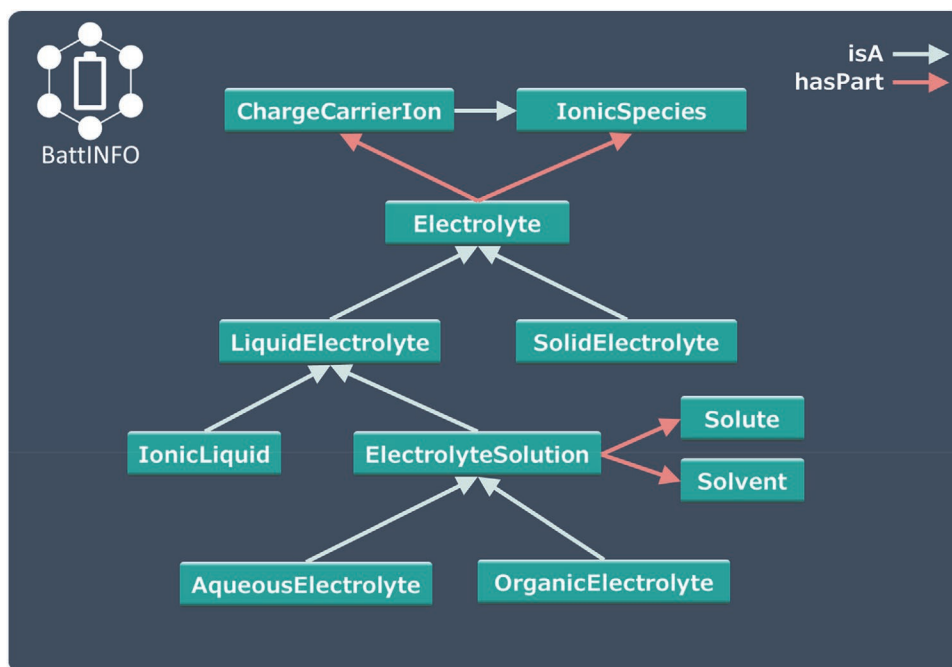
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
ENERGY
MATERIALS
www.advenergymat.de

Figure 6. Overview of some electrolyte relations in BattINFO.

be equal to the sum of negative charges to maintain the electroneutrality of the whole electrolyte. For an electrochemical cell to function, at least one type of the mobile ions in the electrolyte should participate as an electroactive substance electrochemical reaction at the electrode. For example, in the case of Li-ion battery electrolytes, the charge carrier ion is Li^+ and the counter ion is usually PF_6^-.

4.1.3. Electrochemical Cell

Definitions for an electrochemical cell usually describe it in terms of an assembly of some required parts, as shown by the comparison in Table 3. For example, IUPAC defines an electrochemical cell as a system that consists of at least two electron conductors (electrodes) in contact with ionic conductors (electrolytes);[56] Electrochemical Systems defines it as a system containing two electrodes that allow transport of electrons, separated by an electrolyte that allows movement of ions but blocks movement of electrons.[67] Although they seem quite similar, both definitions have different benefits and shortcomings. On the one hand, the IUPAC version is preferable because it highlights that the cell must contain at least two electrodes, while Electrochemical Systems implies that it contains exactly two electrodes. Although most electrochemical cells have two electrodes, there do exist three-electrode electrochemical cells in which a third electrode is included to act as a reference to for electric potential measurement. On the other hand, the Electrochemical Systems definition is preferable because it states that the electrodes must be separated (i.e., not in contact with each other).

The requirement that the electrodes in an electrochemical cell be separated is potentially a source of a semantic inconsistency, which should be addressed. First, among the battery community, the physical space between the electrodes is often informally referred to as the "separator." The same word is also used to describe a material (e.g., a porous polymer film) that is soaked with electrolyte and placed between the electrodes to prevent an internal short-circuit. There must be a distinction between the region called a separator and the material called a separator. Second, the implication inherent in the word "separator" as a description the domain between the electrodes is incomplete. It conveys the idea that the electrodes must not touch, but it does not convey the idea that the domain must also shuttle ions between the electrodes. The term salt bridge is also sometimes used to describe this domain. A salt bridge is defined by IUPAC as the means of making an electrolytic connection between two half cells without inducing a significant liquid junction potential.[56] But again, this introduces a semantic inconsistency as there is also a specific electrochemical device called a salt bridge. Furthermore, the term salt is misleading because not all ion transport is achieved using dissolved salts. To circumvent these challenges, BattINFO designates the domain forming an ionic connection between two electrodes (or half cells) as an ion bridge. The word "bridge" implies separation between two places and the word "ion" implies that ions travel across the bridge. A "separator" can then be described as a component in a particular type of "ion bridge." Both "ion bridge" and "separator" retain unique IRIs for unambiguous referencing.

The ion bridge fulfils two essential roles in the cell: i) It provides physical separation between the two half cells (thus preventing an internal short circuit) and ii) it provides a means to conduct charge carrier ions between the half cells. An ion bridge must contain some electrolyte material and may also contain an electrochemically inert component to provide structural stability or physical separation.

4.1.4. Battery

The battery domain ontology in BattINFO builds on electrochemical knowledge to express information about the whole battery. Even the top class, Battery, can be the source of some semantic inconsistencies between different fields.

As mentioned above, a cell manufacturer might use the word "battery" to refer to a battery cell, while an EV manufacturer could use the same word to mean a battery pack. While humans might easily understand the difference from the context, machines are not able to make that distinction. To address this issue, BattINFO defines both BatteryPack and BatteryCell as a subclass of Battery, with the relationship noting that a BatteryPack hasPart some BatteryCell and BatteryCell hasPart some ElectrochemicalCell, as shown in Figure 3.

There are a few ways to classify a type of battery cell. It can be classified according to its ability to be recharged (i.e., primary, secondary), the battery chemistry (e.g., Li-ion, Zn-air, etc.), or the format of the cell (e.g., cylindrical, pouch, etc.). BattINFO offers a flexible set of battery classification descriptors that are in alignment with common battery standards (e.g., IEC 60 086).

4.1.5. Measurements, Properties, and Material Relations

Battery data expresses information describing some observable properties of a battery obtained from a real or simulated measurement. For example, an engineer might generate data about a specific battery cell using a cycler to measure the nominal capacity. The nominal capacity is a property of the battery that can be expressed in terms of a physical quantity (i.e., electric charge with SI base units A·s). A battery domain ontology should be able to describe the measurement process, the physical quantities obtained from the measurement, and how these relate to object properties.

A measurement in EMMO is a semiosis process that results in a quantitative comparison of a property of an object with a standard reference. Semiosis is a process in which an interpreter creates a sign to represent some object. EMMO contains terms for expressing the fundamental aspects of a measurement as a process involving some measurement system and generating a measurement result. There is an ongoing effort to create an EMMO domain ontology for characterization based on the CHADA template. BattINFO intends to support the development of this ontology and import it when completed. In the meantime, development in BattINFO is focused on creating classes to ontologize battery measurement equipment and the resulting physical quantities and properties.

Measurements generating battery-related data can come from a wide variety of instruments, from common laboratory instruments like galvanostats, potentiostats, and cyclers, to equipment in Gigafactory production lines or large-scale nation research infrastructure like synchrotrons. Ontologizing the whole of this diverse field necessitates collaborations with other domain ontologies. In its early development, BattINFO is focused on ontologizing measuring instruments needed to support the research and industry infrastructure used in the BIG-MAP consortium. Battery measuring instrument classes in BattINFO define the roles and properties of the instruments in a measurement process, from which individual instances of the instruments can be declared in an application ontology. Within the context of a measurement, these instruments interact with some sample to produce a measurement result, which must be quantified. An overview of relations ontologizing the measurement process in BattINFO is shown in **Figure 7**.
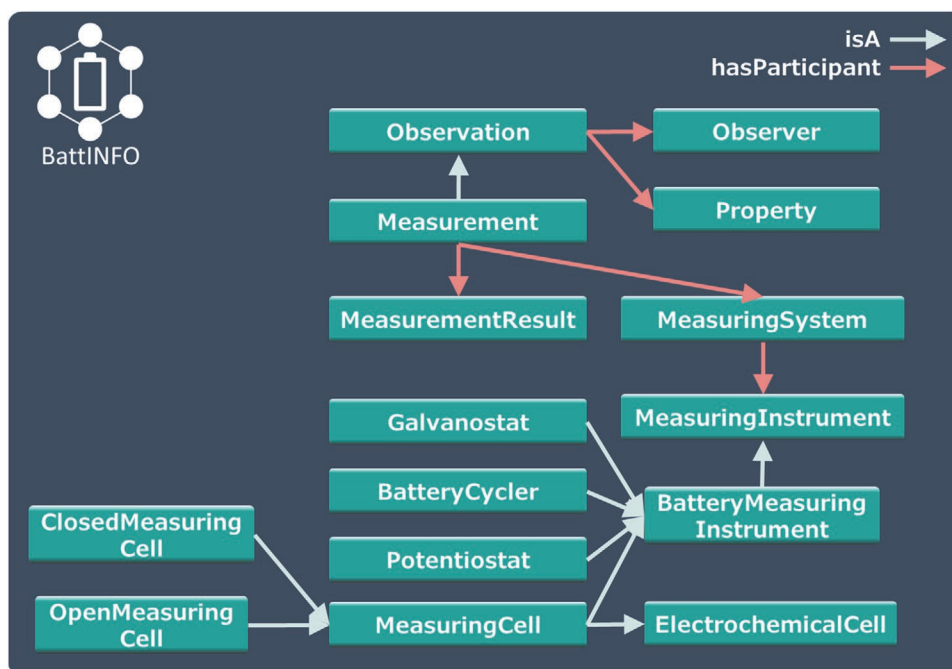

Figure 7. Overview of relations ontologizing the measurement process in BattINFO. Arrows point from the subject to the object.

Table 4. Overview of selected quantities used in the battery domain, as defined in BattINFO; A more extensive list and a guide to interpreting the physical dimensions are provided in Tables S4 and S5, Supporting Information.

Class	Relations	Elucidation	Physical Dimension	Source
ChargeNumber	isA ElectrochemicalKineticQuantity	Number of electrons transferred in a charge transfer reaction between an electrode and a single entity (ion, radical-ion, or molecule) of an electroactive substance, whose identity must be specified.	T0 L0 M0 I0 Θ0 N0 J0	[56]
ChargeTransferCoefficient	isA ElectrochemicalKineticQuantity	The fraction of the electrostatic potential energy affecting the reduction rate in an electrode reaction, with the remaining fraction affecting the corresponding oxidation rate.	T0 L0 M0 I0 Θ0 N0 J0	[73]
ElectrodePotential	isA ElectricPotential	Electric potential at an electrode, reported as the difference in potential relative to a reference electrode.	T-3 L+2 M+1 I-1 Θ0 N0 J0	[56]
ElectrodeSurfaceArea	isA ElectrochemicalQuantity	Area of electrode—solution interface.	T0 L+2 M0 I0 Θ0 N0 J0	[56]
EquilibriumElectrodePotential	isA OpenCircuitPotential	Potential of an electrode when no electric current flows through the cell and all local charge transfer equilibria across phase boundaries that are represented in the cell diagram (except at possible electrolyte-electrolyte junctions) and local chemical equilibria are established.	T-3 L+2 M+1 I-1 Θ0 N0 J0	[56]
ExchangeCurrent	isA ElectricCurrent	The common value (i0) of the anodic and cathodic partial currents when the reaction is at equilibrium.	T0 L0 M0 I+1 Θ0 N0 J0	[57]
FaradaicCurrent	isA ElectricCurrent	Electric current that results from the electrooxidation or electroreduction of an electroactive substance.	T0 L0 M0 I+1 Θ0 N0 J0	[56]
OpenCircuitPotential	isA ElectricPotential	Electrode potential of working electrode relative to the reference electrode when no potential or electric current is being applied to the electrochemical cell.	T-3 L+2 M+1 I-1 Θ0 N0 J0	[56]
Overpotential	isA ElectricPotential	Electrode potential (E) minus the equilibrium electrode potential (Eeq) of an electrochemical reaction.	T-3 L+2 M+1 I-1 Θ0 N0 J0	[56]
SurfaceOverpotential	isA Overpotential	The potential of a working electrode relative to a reference electrode of the same kinds placed in the solution adjacent to the surface of the working electrode (just outside the double layer).	T-3 L+2 M+1 I-1 Θ0 N0 J0	[67]
TransportNumber	isA ElectrochemicalTransportQuantity	Quotient of the current carried by an ionic component and the total current.	T0 L0 M0 I0 Θ0 N0 J0	[57]

Table 4 presents an overview of selected quantities used in the battery domain, listing the name of the class, its relations, elucidation, and physical dimension used in BattINFO. The information needed to ontologize these quantities is taken mostly from the IUPAC recommendations on Terminology for electrochemical methods of analysis[56] and standards for Quantities, Units and Symbols in Physical Chemistry.[72,73] A more extensive list is available in Table S4, Supporting Information. Ontologizing quantities used in electrochemical and battery measurements allows the resulting data to be tagged with appropriate labels. Additional insight can be obtained by describing how quantities are related to each other through equations.

A material relation in EMMO is an equation that stands for a physical assumption specific to a material and provides an expression for a "physics_quantity" (the dependent variable) as function of other variables, physics_quantity or data (independent variables).[74] Equations that form the foundation of electrochemical and battery analysis can then be ontologized as material relations, using the hasSpatialDirectPart relationship to form links between quantities. For example, consider two equations often used in electrochemistry: The Nernst equation and the Butler–Volmer equation. The Nernst equation is an expression derived from the law of mass action that describes the effects of concentration and temperature variations on the equilibrium potential of an electrochemical reaction:

$$E_{eq} = E^{\circ} - \frac{RT}{nF}\ln(Q) \tag{1}$$

where E_{eq} is the equilibrium electrode potential, E° is the standard electrode potential, R is the molar gas constant, T is the temperature, n is the charge number, F is the Faraday constant, and Q is the reaction quotient.

The Butler–Volmer equation, sometimes also called the Butler–Volmer approximation or the Erdy–Gruz–Volmer equation, is a phenomenological model for electrode kinetics, describing the relation between the electrode current from an electrochemical charge-transfer reaction and the surface over-potential of the electrode:

$$i = i_0\left(\exp\left(\frac{\alpha nF}{RT}\eta_{surf}\right) - \exp\left(-\frac{(1-\alpha)nF}{RT}\eta_{surf}\right)\right) \tag{2}$$

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

ADVANCED
ENERGY
MATERIALS
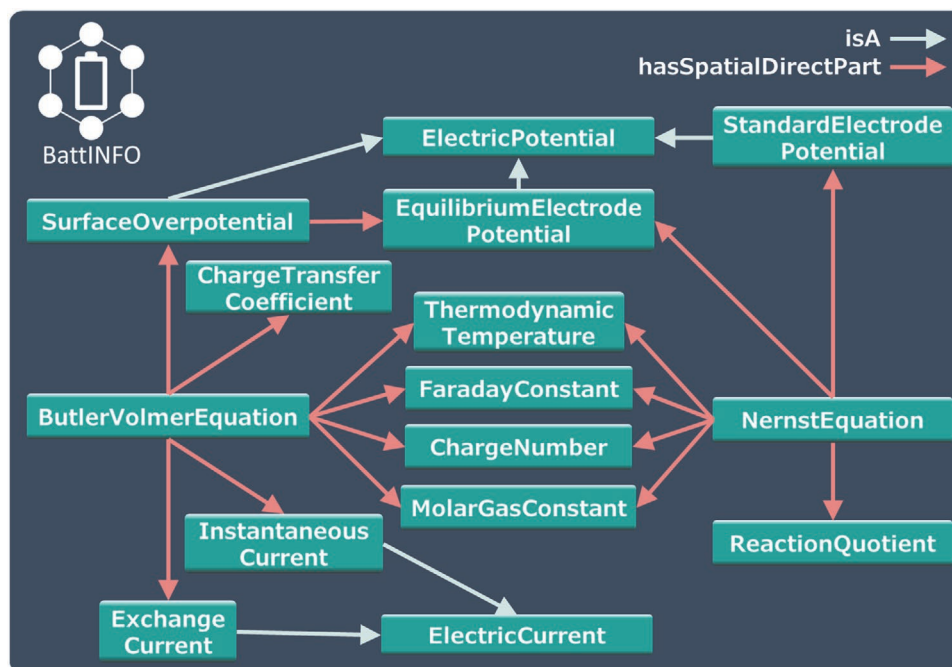
www.advenergymat.de

Figure 8. Simplified overview of quantities coupling the electrochemical relations in the Butler–Volmer equation and the Nernst equation.

where i is the electrode current, i_0 is the exchange current, α is the charge transfer coefficient, and η_{surf} is the surface overpotential:

$$\eta_{surf} = E_{elde} - E_{elyte} - E_{eq} \tag{3}$$

Figure 8 shows a simplified overview of the relationships used to ontologize the Nernst and Butler–Volmer equations. Including these descriptions in BattINFO allows machines to reason about the quantities involved. For example, even though the quantity electrode equilibrium potential does not appear explicitly in the Butler–Volmer equation, using relationships over the surface overpotential clarifies that this quantity is a point of coupling with the Nernst equation. Although this is clear to humans working in the field of electrochemistry, expressing these relations in a machine-readable format paves the way for future work in computational battery research.

4.1.6. Battery Interface Ontology Outlook

BattINFO is a living repository of battery knowledge, and it will continue to grow and develop through both the work of the BIG-MAP consortium and its engagement with the larger battery community. As of August 2021, BattINFO contains 271 terms, and many of these (205) are factored out into an "electrochemistry" branch for easier distinction between these domains if necessary. Furthermore, around 30 more generic terms that might be found in the EMMO ontology itself at a later stage are factored out into a sub-branch.

Readers are encouraged to use the free ontology editor Protégé to interact directly with BattINFO.[34] A Semantic MediaWiki is also in development for users to view BattINFO in a

web browser. The BattINFO repository is available to the public on GitHub. The repository facilitates public comment and is intended to support open development of the ontology. The first stable release of BattINFO is planned for February 2022.[65]

4.2. Battery Value Chain Ontology

The BVCO is a sister ontology of BattINFO, which aims to ontologize essential parts of the battery value chain.[75] BVCO is dedicated to the higher-level process chains for material processing and manufacturing. It imports BattINFO and applies the basic definition of the battery as a system made in that ontology. The BVCO also interacts directly with the General Process Ontology (GPO),[76] which describes terms common across different process engineering domains. The BVCO is being developed under the coordination of Fraunhofer ISC as part of both EU and national research projects. It is hosted in a public repository on GitLab[75] and available for use under a creative commons CC-BY-4.0 license.

To define the battery value chain, BVCO expands the EMMO process definition to include concepts such as, inputs and outputs as well as predecessor and successor. Furthermore, where EMMO limits process participants to matter, BVCO introduces the use of energy and information as process participants. Thus, the processes in BVCO are holistic perspective elements that transform inputs/educts (matter, energy, information) into output/products (matter, energy, information) with the help of tools (devices, algorithms). They can be decomposed into sub-processes and have predecessor and successor processes. Information as process output plays an important role especially for measurement processes while energy is of central importance for the LCA evaluation of process chain.
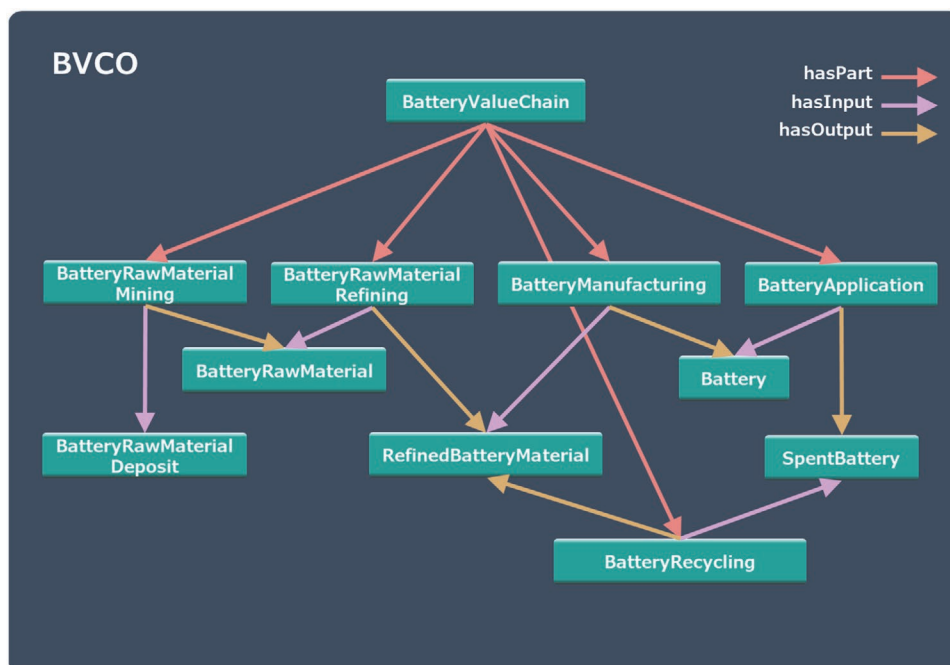
Figure 9. Overview of the battery value chain definition. Arrows point from the subject to the object.

Figure 9 shows an overview of the process steps of the battery value chain, as defined in the BVCO. The value chain begins with the mining of battery raw materials such as, lithium and cobalt ores. The battery raw material is then passed to the battery raw material refining process, in which it is refined and purified for use in battery manufacturing. The manufactured batteries are used in some application, such as electric vehicles.

At the end of their lifespan, the spent battery is sent to the battery recycling process to recover raw materials, refined materials or even components, and the cycle begins again. Each of these high-level processes consists of many sub-processes.

Figure 10 shows a simplified overview of some of the battery manufacturing sub-processes, which are aligned with the BattINFO definition of a battery. The essential steps of battery
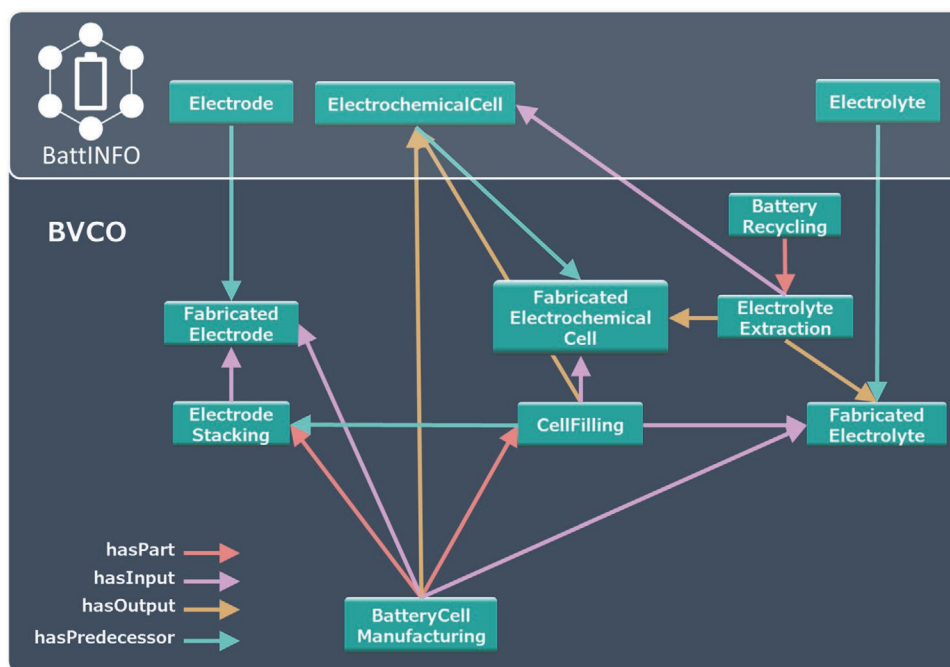

Figure 10. BVCO-BattINFO alignment.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**

www.advenergymat.de

manufacturing, common to all chemistries and cell types, are the production of the base components: Electrodes, electrolyte, and separator. BVCO distinguishes fabricated components from the components themselves. For example, according to the definition of an electrode in BattINFO, the active electrochemical material in the electrode must have contact with some electrolyte. Within the scope of the manufacturing process, the manufacturing object that is intended to be used as an electrode is known as a "fabricated electrode." It only becomes an electrode after the electrolyte is injected during the cell filling step.

For the manufacturing of simple battery cells, only an enclosure is added to the cells made from these components (compare simple cylindrical consumer batteries). In more complex systems, such as an EV battery, the cells are bundled into modules and packs and assembled into a battery system. In addition to the actual cells and an enclosure, this then also includes supporting components such as, electronic battery management systems or cooling systems.

In summary, BVCO defines physical material parts as the result of manufacturing processes (fabricated parts). BattINFO defines functional parts within the functional electrochemical system (functional parts with a role). At certain points in time there is a transition between the two ontologies and concepts, such as during manufacturing when the cell is filled with electrolyte or when the electrolyte is removed during battery recycling (see Figure 10).

5. Applications of a Battery Ontology

The fundamental application of an ontology is to improve communication between humans and/or computers.[31,77] Before discussing the details of specific applications of a battery ontology, we first highlight some prominent examples of ontological applications in other fields to provide context and guidance. These examples include: i) Comparison to the Gene Ontology resource, ii) multi-scale modelling workflows, and iii) digital twins.

Applied ontology has been used to great effect in the life sciences, with the most well-known example being the Gene Ontology (GO) resource. The GO resource provides a computational representation of current scientific knowledge about the functions of genes from many different organisms.[27] The GO offers two primary resources: i) The GO itself, which provides a logical structure describing the full complexity of biology, and ii) the body of GO annotations which provide traceable, evidence-based statements relating a specific gene product to specific ontology terms to describe its normal biological role. GO annotations are statements about how a gene functions at the molecular level, where in the cell it functions, and what biological processes it helps carry out. A battery ontology like BattINFO could provide similar resources. The ontology itself offers a logical structure describing electrochemistry, which could be of great utility to research and development in a variety of fields beyond batteries (e.g., fuel cells, electrolyzers, corrosion, etc.). Creating and compiling BattINFO annotations could, for example, relate specific molecules to the roles they play in battery systems.

Going beyond the life sciences, the use of ontologies in technical and software domains is also growing. One area which has received recent interest is ontology-based multi-scale modelling workflows. The concept of multi-scale modelling is applicable to a variety of systems, whose performance is determined by processes occurring on different time and length scales. While there exist well-developed methods to model individual processes on their own characteristic scales, establishing links between the scales has proved difficult.[78] An interoperability framework for computational tools spanning different time and length scales was recently proposed by Mir et al. and demonstrated by modelling steel corrosion in concrete.[79] In that example, the authors linked DFT simulations of material properties performed in VASP to continuum-scale simulations of the corrosion process performed in COMSOL Multiphysics by passing data inputs and outputs in JSON files and storing the results in a common HDF5 file format. Multi-scale modelling is an essential tool in battery development, and the inefficiency of coupling scales is one of the main challenges slowing its use. A battery ontology that builds on these activities ontologizing modelling workflows could have great impact facilitating not only multi-scale couplings, but also linking to parameter databases and open-source modelling frameworks to reduce the time and cost needed to parameterize models.[80]

Digital twins are another area that have seen recent progress in ontology-driven approaches. A digital twin is a virtual representation of some real component or system. It must contain information about the properties of the real component, models that describe its performance, and sensor data describing the state of the system (often in real time). The vision for smart manufacturing in modern factories is one area that has seen concerted efforts to develop ontology-driven digital twins.[81–83] Implementing standards for reliable and efficient interoperability of data and tools is essential to the functionality of digital twins. Furthermore, the meaning of the data itself must be correctly understood. Ontologies help make this happen. In the scientific literature, there are many recent examples demonstrating the development of ontology-based digital twins for manufacturing and other cyber physical systems.[84–86] Ontology-based approaches are also integrated into commercial digital twin platforms, such as Microsoft Azure which implements customized ontologies in its JSON-LD based Digital Twin Definition Language.

The rest of this section discusses potential applications of a battery ontology, which can be grouped into three broad categories: i) Neutral authoring, ii) concept-based search, and iii) common access to information.

5.1. Neutral Authoring

The concept of neutral authoring is to describe an artifact (e.g., a dataset, software, or other domain ontology) using a single language, such that it can be converted into a different form to be used by multiple target systems.[77]

Data stored and exchanged following a neutral format outlined by a battery ontology improve compliance with FAIR data principles. The ability to reuse data beyond its initial scope benefits i) academic researchers, by maximizing the visibility

of their results, ii) data scientists, by unlocking access to new sources of data, and iii) industries, by minimizing redundant activities and so the costs of generating data. However, in large research projects or battery Gigafactory supply chains, the absence of commonly defined data descriptors and exchange formats are large barriers in the development of an interoperable data infrastructure. As an example, a computational chemist might state that an infrared spectrum of a positive electrode active material must include some measure of sample purity, to assess deviations between simulated and measured spectra. For an experimental electrochemist, the same dataset must specify the sample conditions during measurement (e.g., operando, post-mortem, etc.) to understand whether dynamic cycling processes are expected to influence the spectrum. Besides requiring different data descriptors, researchers likely use data formats that are incompatible with each other.

Ontologies may solve these issues. In essence, ontologists and infrastructure developers decide on using a common, neutral language to locate and exchange data across the domain. Developers then build parsers and application interfaces complying with the agreed language. Last, users operate the developed tools to store and query data. A domain-neutral language may serve to automate not only the translation of datasets into different formats, but also the generation of parsers and application interfaces to support new data flows.[77] Such an infrastructure might, for instance, query a relevant microscope image and autonomously parse it from its proprietary binary format into a 2D array of contrasts for numerical analysis, or into a graph-modelling format for training graph neural networks, or print it as a raster graphics (e.g., jpeg) for visualizing in a document.

Ontologies can also support the development of software tools for the battery domain. An intelligent code design is imperative for engineering interoperable, flexible software. To this end, code written in a specific programming language can be structured mirroring the coherent hierarchy of concepts encoded in the ontology, resulting in a codebase that naturally inherits a logical structure and renders maintainable and understandable software. More importantly, the one-to-one correspondence between code and ontological concepts enables translating the code into other languages via concept-object mappings.[87] For instance, an ontology-compliant modelling software might be written in a high-level, user-friendly language as Python, and be translated to a lower-level but highly efficient language such as C to benefit from improvements in computational performance. When software is developed within an interoperable framework, developers gain the freedom of coding in their preferred language while enjoying access to heterogeneous resources from peers.

The Global Battery Alliance (GBA) has called for immediate action to enable the exchange of battery data using a common battery passport.[3] The GBA envisions the battery passport as a digital asset that conveys all the necessary environmental, social, governance, and lifecycle requirements for the battery to ensure compliance with laws and regulations. However, creating a standardized digital framework for sharing information about specific individual batteries has great potential beyond simply tracking sustainability targets and regulatory compliance. Different approaches for a battery passport have been proposed focusing on sharing data across the battery supply chain[88] or monitoring performance data.[89] A battery passport requires, at a minimum, a standardized language for describing data based on a shared concept of a battery. However, to reach its full potential, a battery passport must be based on a language that is machine-readable and enhanced with semantic descriptions.

5.2. Concept-Based Search

An ontology-based database (OBDB) annotates datasets with semantic concepts that can be queried and explored with powerful computational tools.[90] Take as an example a tomography image of a $LiNiO_2$ composite electrode stored in a repository. Using a conventional database, the image dataset would be tagged with identifiers such as "X-ray tomography," "$LiNiO_2$," while in an OBDB it would be associated with ontological classes like "tomographic measurement" or "positive electrode." A human user querying for "cathode particle volume expansion" would struggle finding the image from the conventional database as the query does not contain any of the tags. In contrast, an ODBD-aware search engine would identify the image as valid query result since it would recognize that the "tomographic measurement" class is associated to the determination of particle volumes, and the "positive electrode" concept is described as synonym of cathode. Synonyms, parenthood, and parthood are some of the relationships that make concept-based search possible, and they are all formally encoded in the ontology.

Machines also benefit from the ability of accessing highly heterogeneous datasets using flexible queries. An OBDB could support workflows where a simulation accesses datasets during runtime. For instance, a continuum simulation of a Li-ion cell could automatically request ab initio datasets from an OBDB to parametrize the transport equations (e.g., diffusion coefficients from an ab initio molecular dynamics simulation), compute the results and compare to experimental voltage curves also stored in the ODBD. Such workflow can be designed to autonomously explore large dataset spaces with minimal human intervention. Likewise, software may also be indexed with ontological tags to facilitate storage, query, access, and interoperability with third-party applications.

5.3. Common Access to Information

Ontologies may be applied to facilitate common access to heterogeneous sources of information among multiple targets (humans or machines).[77]

The expansion of Li-ion battery Gigafactories around the world and the billions of new cells entering the market are the source of some concern regarding both regulatory compliance and recycling.[88] Furthermore, battery passports are intended to be digital information assets accompanying a battery through its entire lifetime. Both Gigafactories and battery passports require some universal way to describe the manufacturing information, materials, or performance history of a given battery cell. Ontologies meet these needs by powering formal tools to encode, consult, and exchange information.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**

www.advenergymat.de

From the perspective of human communication, a battery ontology such as, BattINFO includes a canonical conceptualization of a battery as an electrochemical energy storage device that human actors can consult for definitions, relevant observables (e.g., electrochemical properties), and relationships among concepts within the domain. In this application, the knowledge encoded in the ontology can be represented as digital glossaries using, for instance, semantic wikis, where users browse concepts, definitions, etc., within a Wikipedia-like interface.[91] A unified description of a battery cell, accessible through user-friendly tools, encourages scientists to adopt a common and citable vocabulary for reporting research work and arguing scientific findings, which altogether improve the transmission of knowledge and the impact of their research work.

Common access to data is equally crucial in research and industrial applications. Data annotated with structured, well-defined ontological concepts help both humans and computers to understand the provenance of data, its spatiotemporal domain, and the observable being characterized by the data-generating process (e.g., an experiment, a simulation, etc.). In addition, as mentioned above, neutrally-authored and ontological indexed data ensures that multiple actors can access and reuse the same resources independently of how and why these resources were generated. Ontologies also facilitate common access to code. Software tagged with ontological concepts not only improves accessibility as discussed above, but also enriches code documentation. In this way, developers and end-users establish a common language to facilitate communication and development.

Ontologies may also be used to power digital research infrastructures. In particular, ontologies can leverage the development of multiple standardized application programming interfaces (APIs) to build a constellation of shared services.[77] Just as web microservices manage enormous fluxes of data in large websites,[92] research-oriented shared services may orchestrate the interchange and analysis of scientific data. Consider as an example a user searching for the Li^+ coordination number in a specific electrolyte. In a shared services infrastructure, the user would submit a query to a client API that in turn coordinates calls to multiple services: A search via a OBDB-powered engine, a service retrieving relevant resources (e.g., the Raman spectra and DFT-calculated molecular structures of the electrolyte), a verification of whether the data is raw or processed, a service requesting further data processing scripts to estimate the coordination number (e.g., from the spectrum and the DFT-relaxed structures), and finally a service to retrieve the result as a structured table. Each service is accessed through its associated API; hence, for such infrastructure to work, the API constellation must be aware of the relations among the query, the data resources and the relevant software. Ontologies not only provide an explicit encoding of these conceptual relations, but also may automate the challenging process of developing API hives with the help of interface generators.[77]

From the perspective of machines as targets, ontologies may unlock access to large amounts of research data to train artificial intelligence (AI) algorithms. While recent milestones in the development AI have taken place in areas with high or unlimited data availability, (e.g., speech processing, image processing, and game AI), the pace in research-oriented AI applications is severely limited by the enormous cost of acquiring scientific data, especially in battery research.

Text-mining is receiving growing interest as a means of compiling battery databases from natural language information published in the scientific literature.[93–95] Tools such as, ChemicalTagger[96] and ChemDataExtractor[97] use natural language processing to automatically extract information about chemical entities, properties, and relationships directly from research articles. In this context, a battery domain ontology may serve to define a common format supporting the ingestion and interchange of heterogeneous datasets, compiled from mining the battery research literature.

Alternatively, the demand for experimental data to train AI algorithms can be significantly reduced in multiple ways. For one, experimental data can be complemented with virtual data generated from simulation models, which are comparably more scalable and easier to automate.[98] However, the availability of simulation models varies widely across applications. While powerful ab initio methods are readily available to calculate macroscopic properties of metals, the modelling of polymerization reactions is still very challenging. Even more difficult is the simulation of complex systems such as a battery, in which various chemical and physical effects couple over several scales. In these areas in particular, the inclusion of explicit knowledge is therefore additionally required to reduce data requirements.

This enables AI to include explicit relationships between data in addition to data (hybrid AI). Explicit relationships can be, for example, unambiguous relationships such as chemical constituents or physical laws, but also weaker forms such as an assumed but unquantifiable relationship between parameters. However, this requires that the explicit knowledge is expressed (formalized) in a machine-readable form. This is exactly what ontologies can do by transferring the almost unlimited complexity of human knowledge into a strict description language without loss of information. At the same time, they enforce a standardization, which is necessary for the multidisciplinary consolidation of explicit knowledge. The knowledge contained in BattINFO and BVCO about the internal relationships in the battery and the external relationships in its manufacturing can therefore help to significantly reduce the data requirements in future AI applications.

6. Challenges

The creation of ontologies for linked battery data in the Semantic Web is the subject of intense research and development around the world. Although this initiative is making fast progress and building on the current momentum of data-driven approaches across technical disciplines, it does face some challenges. These challenges deal primarily with the vast scope of knowledge and heterogeneous data required in the battery domain. Specific challenges include: i) Creating and integrating well-developed domain ontologies, ii) engaging the battery community in ontology development and application, and iii) applying ontologies in existing digital tools for processing battery data.

BattINFO begins by creating a domain ontology of electrochemistry. The proposed electrochemistry ontology is the

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
ENERGY
MATERIALS
www.advenergymat.de

most advanced currently available, but it is impossible to fully separate electrochemistry from other domains such as physical chemistry, thermodynamics, kinetics, and transport phenomena. In the absence of a well-developed and maintained domain ontology ecosystem, the task of ontologizing a single domain can quickly creep beyond its intended scope. Although relevant domain ontologies exist and others are in development, there are currently too few to fully cover the needs of diverse technical fields like battery research. This challenge can be addressed through the establishment of clear standards for domain ontology generation overseen by an active and engaged coordinating body.

The success of an ontology depends on developing an active base of contributors and users. It is essential that there exist quick and easy pathways for users to access, understand, and apply the ontology in their work. Researchers and engineers worldwide must be able to share their research data and findings in semantically readable formats without being experts in data science themselves. The BIG-MAP project will provide training materials to introduce new users to the ontology and demonstrate the creation of BattINFO compliant metadata. One example could be the use of BattINFO IRIs as property names in JSON-LD metadata. Encouraging the long-term adoption of the ontology in battery workflows requires a two-tiered solution: i) Engaging the community in the development of the ontologies and ii) creating simple user-interfaces for generating and sharing metadata.

The task of bringing together a volunteer community to create and manage knowledge content is a challenge, but there exist guiding precedents to achieve it. Wikis are hypertext publications that can be collaboratively created and edited by their users in a web browser. Semantic MediaWiki is a type of wiki enables users to annotate content with explicit, machine-readable information.[99] Semantic MediaWikis offer a simple Wikipedia-like interface for users to browse the ontology, participate in discussions, and submit requests for modifications.[100,101] A Semantic MediaWiki is currently in development for the battery domain.[102] Engaging the community through benefits the ontology development by ensuring that it is sufficient to describe the relevant data. On the other hand, the wiki also provides users with a consolidated source of battery knowledge that can support their research or design activities. However, developing robust ontologies and engaging with the community is only half the battle. To be successful, the ontologies must be actively used to share battery data.

The burden of time and effort needed to share data in a standardized format and generate the appropriate metadata is one of the main hurdles discouraging the uptake of data into the semantic web. Easy-to-navigate graphical user interfaces and semi-automated conversion tools are needed. The use of electronic lab notebooks coupled to semantic data repositories can help address this challenge. There are a few notable initiatives to develop digital infrastructure for handling battery data, including (but not limited to) Kadi4mat, the Battery Archive, battery evaluation and early prediction (BEEP), and Galvanalyser.

The Kadi4mat initiative has developed digital research data infrastructure for materials science to enable structured data storage and exchange with documented and reproducible data analysis and visualization.[103] Kadi4mat combines an electronic lab notebook interface and data repository with tools for efficient searching and retrieval of data. Furthermore, simple workflows can be defined with a web-based node editor and be run locally. Although Kadi4mat was ostensibly developed for materials science, early use cases have focused on research data from the battery domain.[104] Kadi4mat is a recent development, with the first beta release in 2020. In the long-term, it aims to offer well-defined semantics by integrating domain ontologies based on the EMMO.

In the battery community, there are similar initiatives to support the easy management and exchange of battery data between devices and groups. The Battery Archive is a web-based repository supported by the United States Department of Energy for easy visualization, analysis, and comparison of battery data across institutions.[105–107] Battery data generated by different entities can be submitted to the site, where it is converted into a standard format to allow for easy cross-comparison.[108]

The BEEP tool is an open-source python package to manage high-throughput battery cycling data.[109] BEEP is designed to automate the process of organizing, parsing, and structuring large battery datasets to support exchange between research groups and machine-learning workflows. Similarly, researchers at Oxford are developing a solution called Galvanalyser to support standardization and exchange of data for different battery testing equipment.[110] Galvanalyser is a database for collecting and collating different types of battery test data, with a focus on data acquisition, parsing, and storage. The concept is to automatically collect raw data generated by battery test equipment, parse it into a common format, and store it in a single PostgreSQL database. It is currently undergoing early trials and not yet available to the wider battery community.

Although each of these tools approaches the challenge of battery data management in a different way, they all serve the common goals of reducing the burden for researchers to share their data and ensuring that the data adheres to common formats and metadata standards. Efforts are underway to align the further development of battery data management tools into at least one open-source suite of interoperable software tools that align experimental protocols and data formats.[111,112]

7. Summary

The future of the battery industry depends on data. Data drives the discovery of new battery materials, it optimizes the links between manufacturing and performance, it gives engineers critical insight into the health and lifetime of their products, and it allows recyclers to efficiently recover raw materials.

Around the world, battery research labs and Gigafactories are generating an unprecedented wealth of battery data. However, the battery community has so far only utilized a fraction of its full potential. The problem is that data is often isolated and insufficiently annotated, so that it can only be used at its source for a limited purpose. To extract the full value of data, it must be interoperable. Data generated by different devices made by different manufacturers should be able to be seamlessly exchanged and understood by different computers on a network. Realizing this goal requires the development of a common machine-readable language for annotating battery data.

A common language is needed to describe battery data, so that different groups or devices use the same unambiguous vocabulary when referring to a given concept. But this need extends beyond simply aligning terminology. In today's research and industrial landscape, there is a growing need for computers to interact directly with data. While humans can infer information and relationships about data from the context, machines need explicit instructions to understand data and its meaning. Therefore, the common language we develop must not only be human-readable, but also provide machine-readable, semantically rich information about the data itself and its relationships to other data.

Ontologies offer an excellent tool for creating a common conceptualization of the battery domain and applying it to semantically annotate battery data. The EMMO is a multidisciplinary effort to develop a standard representational framework for applied sciences. It provides a top-level ontology based on physics, analytical philosophy and information and communication technologies to provide a framework for knowledge capture that is consistent with scientific principles and methodologies. EMMO hosts an eco-system of domain ontologies to support interoperability of data in the battery community. There are currently two main initiatives aiming to create fully open EMMO-based battery domain ontologies: the BattINFO and the BVCO.

BattINFO is an ontology of battery cells, materials, and their interfaces developed by the BIG-MAP consortium. It is developed with the goal of creating a chemistry-neutral formalized description of batteries to support data interoperability and artificial intelligence workflows. While BattINFO ontologizes the battery domain from the cell-level and down, the BVCO ontologizes the battery value chain. Considering the process of battery manufacturing from raw materials mining all the way through applications and recycling, the BVCO supports interoperability of data between different actors in the value chain.

Using ontologies to create unified descriptions of battery data has the potential to open the battery field to a new era of open research and development. A battery ontology can support visions for a digital battery passport to share manufacturing information and performance history about a battery across its lifetime. It can distil the deluge of data currently being generated in laboratories, factories, and field applications around the world to artificial intelligence workflows. It can support standardized reporting of research results to help address the reproducibility crisis in scientific publishing. Ultimately, the battery ontologies reviewed in this work are a resource for the community. They provide a well-developed source of battery knowledge and give researchers and engineers the tools they need to get the greatest value from their data.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Keywords

[1] M. S. Whittingham, *Nano Lett.* **2020**, *20*, 8435.

[2] A. A. Tidblad, K. Edström, G. Hernández, I. de Meatza, I. Landa-Medrano, J. J. Biendicho, L. Trilla, M. Buysse, M. Ierides, B. P. Horno, Y. Kotak, H. G. Schweiger, D. Koch, B. S. Kotak, *Energies* **2021**, *14*, 4223.

[3] World Economic Forum, *Global Battery Alliance, A Vision for a Sustainable Battery Value Chain in 2030*, World Economic Forum, Geneva, Switzerland **2019**, https://www3.weforum.org/docs/WEF_A_Vision_for_a_Sustainable_Battery_Value_Chain_in_2030_Report.pdf (accessed: August 2021).

[4] IPCC, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK **2021**.

[5] S. Babinec, E. J. Dufek, D. A. Howey, V. Viswanathan, M. Aykol, D. A. C. Beck, B. Blaiszik, B. R. Chen, G. Crabtree, V. De Angelis, P. Dechent, M. Dubarry, E. E. Eggleton, D. P. Finegan, I. Foster, C. Gopal, P. K. Herring, V. W. Hu, N. H. Paulson, Y. Preger, D. U. Sauer, K. Smith, S. W. Snyder, S. Sripad, T. R. Tanim, L. Teo, Principles of the Battery Data Genome, (Preprint), arXiv:2109.07278, v3, submitted Sept. **2021**.

[6] V. Sulzer, P. Mohtat, A. Aitio, S. Lee, Y. T. Yeh, F. Steinbacher, M. U. Khan, J. W. Lee, J. B. Siegel, A. G. Stefanopoulou, D. A. Howey, *Joule* **2021**, *5*, 1934.

[7] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, A. A. Lee, *Nat. Commun.* **2020**, *11*, 1706.

[8] P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y. H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon, W. C. Chueh, *Nature* **2020**, *578*, 397.

[9] M. Aykol, P. Herring, A. Anapolsky, *Nat. Rev. Mater.* **2020**, *5*, 725.

[10] G. dos Reis, C. Strange, M. Yadav, S. Li, *Energy AI* **2021**, *5*, 100081.

[11] M. D. Wilkinso, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. d. S. Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, et al., *Sci. Data* **2016**, *3*, 1.

[12] B. Mons, C. Neylon, J. Velterop, M. Dumontier, L. O. B. D. S. Santos, M. D. Wilkinson, *Inf. Serv. Use* **2017**, *37*, 49.

[13] Y.-K. Sun, *ACS Energy Lett.* **2021**, *6*, 2187.

[14] A. K. Stephan, *Joule* **2021**, *5*, 1.

[15] D. Stock, S. Dongmo, J. Janek, D. Schröder, *ACS Energy Lett.* **2019**, *4*, 1287.

[16] J. Li, C. Arbizzani, S. Kjelstrup, J. Xiao, Y. Y. Xia, Y. Yu, Y. Yang, I. Belharouak, T. Zawodzinski, S. T. Myung, R. Raccichini, S. Passerini, *J. Power Sources* **2020**, *452*, 227824.

[17] B. Wu, Y. Yang, D. Liu, C. Niu, M. Gross, L. Seymour, H. Lee, P. M. L. Le, T. D. Vo, Z. D. Deng, E. J. Dufek, M. S. Whittingham, J. Liu, J. Xiao, *J. Electrochem. Soc.* **2020**, *166*, A4141.

[18] S. Chen, C. Niu, H. Lee, Q. Li, L. Yu, W. Xu, J. G. Zhang, E. J. Dufek, M. S. Whittingham, S. Meng, J. Xiao, J. Liu, *Joule* **2019**, *3*, 1094.

[19] M. Dubarry, G. Baure, *Electron* **2020**, *9*, 152.

[20] H. Li, *Joule* **2019**, *3*, 911.

[21] Z. Lin, T. Liu, X. Ai, C. Liang, *Nat. Commun.* **2018**, *9*, 8.

[22] A. J. Bard, G. Inzelt, F. Scholz, in *Electrochemical Dictionary*, 2nd ed., Springer-Verlag, Berlin **2012**.

[23] B. Smith, B. Klagges, B. Henning, P. Grenon, T. Bittner, L. Jansen, P. Heuer, U. Schwarz, I. Johansson, *Applied Ontology An Introduction*, Ontos Verlag, Heusenstamm **2008**.

[24] M. Bennett, K. Baclawski, *Appl. Ontol.* **2017**, *12*, 189.

[25] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, *Nat. Genet.* **2000**, *25*, 25.

[26] The Gene Ontology Consortium, *Nucleic Acids Res.* **2019**, *47*, D330.

[27] Gene Ontology Resource, http://geneontology.org/ (accessed: August 2021).

[28] T. Berners-Lee, R. Cailliau, A. Luotonen, H. F. Nielsen, A. Secret, *Commun. ACM* **1994**, *37*, 76.

[29] T. Berners-Lee, J. Hendler, O. Lassila, *Sci. Am.* **2001**, *284*, 34.

[30] C. Bizer, T. Heath, T. Berners-Lee, in *Semantic Services, Interoperability and Web Applications: Emerging Concepts* (Ed: A. Sheth), **2011**, Ch. 8.

[31] D. Fensel, S. Staab, R. Studer, F. van Harmelen, J. Davies, in *Towards the Semantic Web: Ontology–driven Knowledge Management* (Eds: J. Davies, D. Fensel, F. van Harmelen), John Wiley & Sons Ltd, Chichester, UK **2002**, Ch. 14.

[32] D. L. McGuinness, F. van Harmelen, *OWL Web Ontology Language* **2004**, https://www.w3.org/TR/owl-features/ (accessed: August 2021).

[33] B. C. Grau, I. Horrocks, B. Motik, B. Parsia, P. Patel-Schneider, U. Sattler, *J. Web Semantics* **2008**, *6*, 309.

[34] Protege, https://protege.stanford.edu/ (accessed: August 2021).

[35] Protege Wiki, https://protegewiki.stanford.edu/wiki/Main_Page (accessed: August 2021).

[36] J. B. Lamy, *Artif. Intell. Med.* **2017**, *80*, 11.

[37] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, Z. Wang, *J. Autom. Reasoning* **2014**, *53*, 245.

[38] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, Y. Katz, *J. Web Semantics* **2007**, *5*, 51.

[39] Python API for EMMO, https://github.com/emmo-repo/EMMO-python (accessed: August 2021).

[40] D. Tsarkov, I. Horrocks, in *Automated Reasoning* (Eds: U. Furbach, N. Shankar), Springer, Berlin **2006**.

[41] T. Gruber, in *Encyclopedia of Database Systems* (Eds: L. Liu, M. T. Özsu), Springer New York, New York, **2016**, p. 1.

[42] C. Partridge, A. Mitchell, A. Cook, D. Leal, J. Sullivan, M. West, *A survey of Top-Level Ontologies – to inform the ontological choices for a Foundation Data Model*, CDBB, **2020**, https://doi.org/10.17863/CAM.58311 (accessed: August 2021).

[43] European Materials Modelling Council, *The EMMC Roadmap for Materials Modelling and Digitalisation of the Materials Sciences*, Brussels, Belgium, EMMC ASBL, **2020**, https://zenodo.org/record/4272033#.YaP32_nMKUk (accessed: August 2021).

[44] A. C. Varzi, *Appl. Ontol.* **2005**, *2*, 141.

[45] C. S. Peirce, A. Marquand, C. Ladd-Franklin, O. H. Mitchell, B. I. Gilman, *Studies in logic. By members of the Johns Hopkins university*, Little, Brown, and Company, Boston **1883**.

[46] IUCr, https://www.iucr.org/resources/cif/ddl/ddlm (accessed: August 2021).

[47] IUCr, https://www.iucr.org/resources/cif/dictionaries (accessed: August 2021).

[48] M. T. Horsch, S. Chiacchiera, M. A. Seaton, I. T. Todorov, K. Šindelka, M. Lísa, B. Andreon, E. B. Kaiser, G. Mogni, G. Goldbeck, R. Kunze, G. Summer, A. Fiseni, H. Brüning, P. Schiffels, W. L. Cavalcanti, *KI – Kunstl. Intelligenz* **2020**, *34*, 423.

[49] A. F. de Baas, *What makes a material function? Let me compute the ways: Modelling in H2020 LEIT-NMBP programme materials and nanotechnology projects*, 6th ed., Publications Office of the European Union, Luxembourg **2017**.

[50] European Committee for Standardization, *(CEN), CWA Vol. 17284:2018 Materials modelling – Terminology, classification and metadata*, CEN, Brussels, Belgium **2018**, https://emmc.info/wp-content/uploads/2018/05/CWA_17284.pdf (accessed: August 2021).

[51] *OSMO: Ontology for Simulation, Modelling, and Optimization*, https://zenodo.org/record/5084394#.YS587fozaUk (accessed: August 2021).

[52] N. Romanos, M. Kalogerini, E. P. Koumoulos, A. K. Morozinis, M. Sebastiani, C. Charitidis, *Mater. Today Commun.* **2019**, *20*, 100541.

[53] The Product Types Ontology, http://www.productontology.org/ (accessed: August 2021).

[54] The SEAS Battery Ontology, https://ci.mines-stetienne.fr/seas/BatteryOntology (accessed: August 2021).

[55] Ontology for Sensor Network, https://www.w3.org/2005/Incubator/ssn/ssnx/energy/ssn-energy.html accessed: August 2021).

[56] J. M. Pingarrón, J. Labuda, J. Barek, C. M. A. Brett, M. F. Camões, M. Fojta, D. B. Hibbert, *Pure Appl. Chem.* **2020**, *92*, 641.

[57] IUPAC, *Compendium of Chemical Terminology 2.3.3, IUPAC*, Research Triangle Park, NC, USA, **2021**.

[58] Electrochemistry Dictionary and Encyclopedia, https://knowledge.electrochem.org/ed/dict.htm (accessed: August 2021).

[59] International Electrotechnical Commission, *International electrotechnical vocabulary – Part 114: Electrochemistry*, International Electrotechnical Commission, Geneva, Switzerland **2014**.

[60] *International Electrotechnical Commission, IEC 60050 - International Electrotechnical Vocabulary* **2021**, https://www.electropedia.org/ (accessed: August 2021).

[61] G. Deepak, A. Ayush Kumar, A. Santhanavijayan, N. Prakash, in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, IEEE **2019**.

[62] *International Electrotechnical Commission, IEC60086, 2021*, International Electrotechnical Commission, Geneva, Switzerland **2021**.

[63] IEEE, *IEEE Standard Glossary of Stationary Battery Terminology*, IEEE, New York **2016**.

[64] N. von Drachenfels, F. Cerdas, C. Herrmann, *Proc. CIRP* **2020**, *90*, 683.

[65] Battery Interface Ontology (BattINFO), https://github.com/BIG-MAP/BattINFO (accessed: August 2021).

[66] BIG-MAP, https://www.big-map.eu/ (accessed: August 2021).

[67] J. Newman, K. E. Thmoas-Alyea, *Electrochemical Systems*, John Wiley & Sons, Hoboken, NJ **2004**.

[68] A. J. Bard, L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, John Wiley & Sons, New York **2001**.

[69] D. Linden, T. Reddy, *Handbook of Batteries*, McGraw Hill, New York **2002**.

[70] M. R. Palacin, *Accounts Mater. Res.* **2021**, *2*, 319.

[71] M. Marcinek, J. Syzdek, M. Marczewski, M. Piszcz, L. Niedzicki, M. Kalita, A. Plewa-Marczewska, A. Bitner, P. Wieczorek, T. Trzeciak, M. Kasprzyk, P. Łężak, Z. Z. A. Zalewska, W. Wieczorek, *Solid State Ionics* **2015**, *276*, 107.

[72] E. R. Cohen, T. Cvitaš, J. G. Frey, B. Holmström, K. Kuchitsu, R. Marquardt, I. Mills, F. Pavese, M. Quack, J. Stohner, H. L. Straus, A. J. Thor, *Int. Union Pure Appl. Chem.* **2007**, *233*, 11.

[73] R. Guidelli, R. G. Compton, J. M. Feliu, E. Gileadi, J. Lipkowski, W. Schmickler, S. Trasatti, *Pure Appl. Chem.* **2014**, *86*, 245.

[74] Elementary Multiperspective Material Ontology (EMMO), https://github.com/emmo-repo/ (accessed: August 2021).

[75] BVCO, https://gitlab.cc-asp.fraunhofer.de/ISC-Public/ISC-Digital/ontology/bvco (accessed: August 2021).

[76] GPO, https://gitlab.cc-asp.fraunhofer.de/ISC-Public/ISC-Digital/ontology/gpo (accessed: August 2021).

[77] M. Uschold, R. Jasper, in *Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods* (Eds: V. R. Benjamins, B. Chandrasekaran, A. Gomez-Perez, N. Guarino, M. Uschold), Stockholm **1999**.

[78] A. A. Franco, A. Rucci, D. Brandell, C. Frayret, M. Gaberscek, P. Jankowski, P. Johansson, *Chem. Rev.* **2018**, *119*, 4569.

[79] Z. M. Mir, J. Friis, T. F. Hagelien, I. H. Svenum, I. G. Ringdalen, N. Konchakova, M. L. Zheludkevich, D. Höche, *Modell. Simul. Mater. Sci. Eng.* **2020**, *28*, 025003.

[80] A. A. Wang, S. E. J. O'Kane, F. B. Planella, J. L. Houx, K. O'Regan, M. Zyskin, J. Edge, C. W. Monroe, S. J. Cooper, D. A. Howey, E. Kendrick, J. M. Foster, *arXiv* **2021**, 2110.09879.

[81] M. Meyer, Z. Yu, P. Gulati, A. Delforouzi, J. Roggenbuck, K. Wolf, Ontologies for Digital Twins in Smart Manufacturing, Fraunhofer ACSI, Sankt Augustin, Germany **2020**, http://doi.org/10.13140/RG.2.2.11346.17607 (accessed: August 2021).

[82] Y. Cui, S. Kara, K. C. Chan, *Rob. Comput. Integr. Manuf.* **2020**, *62*, 101861.

[83] E. Järvenpää, N. Siltala, O. Hylli, M. Lanz, *J. Intell. Manuf.* **2019**, *30*, 959.

[84] S. Singh, E. Shehab, N. Higgins, K. Fowler, D. Reynolds, J. A. Erkoyuncu, P. Gadd, *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2020**.

[85] E. Negri, L. Fumagalli, C. Cimino, M. MacChi, *Procedia Manuf.* **2019**, *28*, 201.

[86] J. Bao, D. Guo, J. Li, J. Zhang, *Enterp. Inf. Syst.* **2019**, *13*, 534.

[87] H.-J. Happel, S. Seedorf, *Int. J. Prod. Perform. Manag.* **2005**, *54*, 1.

[88] Y. Bai, N. Muralidharan, Y. K. Sun, S. Passerini, M. S. Whittingham, I. Belharouak, *Mater. Today* **2020**, *41*, 304.

[89] W. Li, N. Sengupta, P. Dechent, D. Howey, A. Annaswamy, D. U. Sauer, *J. Power Sources* **2021**, *506*, 230024.

[90] H. Dehainsala, G. Pierra, L. Bellatreche, in *Advances in Databases: Concepts, Systems and Applications* (Eds: R. Kotagiri, P. R. Krishna, M. Mohania, E. Nantajeewarawat), Springer, Berlin **2007**.

[91] F. Bry, S. Schaffert, D. Vrandečić, K. Weiand, in *Reasoning Web. Semantic Technologies for Advanced Query Answering: 8th International Summer School 2012, Vienna, Austria, September 3–8, 2012. Proceedings* (Eds: T. Eiter, T. Krennwallner), Springer, Berlin **2012**, pp. 329–369.

[92] D. Taibi, V. Lenarduzzi, C. Pahl, *IEEE Cloud Comput.* **2017**, *4*, 22.

[93] H. El-Bousiydi, T. Lombardo, E. N. Primo, M. Duquesnoy, M. Morcrette, P. Johansson, P. Simon, A. Grimaud, A. A. Franco, *Batteries Supercaps* **2021**, *4*, 758.

[94] S. Huang, J. M. Cole, *Sci. Data* **2020**, *7*, 260.

[95] A. Torayev, P. C. M. M. Magusin, C. P. Grey, C. Merlet, A. A. Franco, *J. Phys. Mater.* **2019**, *2*, 044004.

[96] L. Hawizy, D. M. Jessop, N. Adams, P. Murray-Rust, *J. Cheminf.* **2011**, *3*, 17.

[97] M. C. Swain, J. M. Cole, *J. Chem. Inf. Model.* **2016**, *56*, 1894.

[98] M. Dubarry, D. Beck, *J. Power Sources* **2020**, *479*, 228806.

[99] M. Krötzsch, D. Vrandečić, M. Völkel, in *The Semantic Web – ISWC*, Springer, Berlin **2006** pp. 935–942.

[100] D. Vrandecic, M. Krötzsch, in *Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics* (Eds: M. Voelkel, S. Schaffert) **2006**, *206*.

[101] C. Willmes, F. Viehberg, S. E. Lopez, G. Bareth, *Data* **2018**, *3*, 44.

[102] Onto-Wiki, https://onto-wiki.eu/ (accessed: August 2021).

[103] N. Brandt, L. Griem, C. Herrmann, E. Schoof, G. Tosato, Y. Zhao, P. Zschumme, M. Selzer, *Data Sci. J.* **2021**, *20*, 8.

[104] J. Santoki, S. Daubner, D. Schneider, M. Kamlah, B. Nestler, *Modell. Simul. Mater. Sci. Eng.* **2021**, *29*, 065010.

[105] Battery Archive, https://www.batteryarchive.org/ (accessed: August 2021).

[106] Y. Preger, H. M. Barkholtz, A. Fresquez, D. L. Campbell, B. W. Juba, J. Romàn-Kustas, S. R. Ferreira, B. Chalamala, *J. Electrochem. Soc.* **2020**, *167*, 120532.

[107] Battery-lfc/batteryarchive, https://github.com/battery-lcf/batteryarchive (accessed: August 2021).

[108] V. De Angelis, Y. Preger, B. R. Chalamala, (Preprint) ecsarxiv, https://ecsarxiv.org/h7c24/, v3, Submitted: March, **2021**.

[109] P. Herring, C. B. Gopal, M. Aykol, J. H. Montoya, A. Anapolsky, P. M. Attia, W. Gent, J. S. Hummelshøj, L. Hung, H. K. Kwon, P. Moore, D. Schweigert, K. A. Severson, S. Suram, Z. Yang, R. D. Braatz, B. D. Storey, *SoftwareX* **2020**, *11*, 100506.

[110] A. Lewis-Douglas, L. Pitt, D. A. Howey, *arXiv*, 2010.14959, **2020**.

[111] T. Vegge, J. Tarascon, K. Edström, *Adv. Energy Mater.* **2021**, *11*, 2100362.

[112] I. E. Castelli, D. J. Arismendi-Arrieta, A. Bhowmik, I. Cekic-Laskovic, S. Clark, R. Dominko, E. Flores, J. Flowers, K. U. Frederiksen, J. Friis, A. Grimaud, K. V. Hansen, L. J. Hardwick, K. Hermansson, L. Königer, H. Lauritzen, F. L. Cras, H. Li, S. Lyonnard, H. Lorrmann, N. Marzari, L. Niedzicki, G. Pizzi, F. Rahmanian, H. Stein, M. Uhrin, W. Wenzel, M. Winter, C. Wölke, T. Vegge, *Batteries Supercaps* **2021**.

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
ENERGY
MATERIALS**
www.advenergymat.de

Simon Clark studied Mechanical Engineering and Computational Electrochemistry. After working as a design simulation consultant in the aerospace industry, he completed his M.Sc. and Dr. rer. nat. degrees at Ulm University in cooperation with the German Aerospace Center (DLR) and Helmholtz Institute Ulm (HIU). Since 2018, he has worked as a Research Scientist at SINTEF Industry, based in Trondheim, Norway. Dr. Clark's research focuses on the development of model-based design tools to improve the performance of electrochemical devices. He currently leads the development of the Battery Interface Ontology (BattINFO).

Francesca L. Bleken earned her Ph.D. in Chemistry focused on experimental heterogeneous catalysis from the University of Oslo in 2011. She is a research scientist at SINTEF, Oslo, since 2013 applying both periodic and non-periodic QM/DFT-modelling to industrial problems with focus on chemical reactions and catalytic systems. Dr. Bleken has a strong interest in interoperability and ontology work targeted towards joining theoretical and experimental disciplines within chemistry. She is actively involved in both developing and applying ontologies, and is one of the core developers of EMMOntoPy as a complementary python tool for this.

Simon Peter Stier began his higher education in 2011 with the admission of two bachelor courses, Aerospace Computer Science and Technology of Functional Materials, at JMU Würzburg. After graduating in 2014, he continued the consecutive Master's programs and received both the Master's degree in Computer Science and in Functional Materials in 2016. From 2017-2020 he did his PhD at Fraunhofer-ISC in the field of interdisciplinary research of sensor materials, manufacturing and corresponding software applications. He then took over as Head of Digital Transformation at Fraunhofer-ISC, where he has made the transition from paper-based to semantically deeply interlinked research his mission.

Eibar Flores received his BSc. in Physics from University of Panama (2013) and his M.Sc. in materials for energy storage and conversion from the European Erasmus Mundus M.E.S.C. master's program (2015). He completed his Dr.Sc. degree from ETH Zurich (2019) working at the Paul Scherrer Institut in Switzerland. Since 2020, Dr. Flores has worked as a Postdoctoral Researcher at the Technical University of Denmark, where he currently focuses on developing digital tools to accelerate battery research such as ontologies, software and data-driven methods to describe, analyse and model research data, respectively.

Casper Welzel Andersen obtained his PhD in Nanoscience in 2018 at Aarhus University, Aarhus, Denmark. He studied industrial zeolite-based deNOx catalysts in collaboration with Haldor Topsøe A/S using mainly powder X-ray diffraction. Since 2018, Dr. Andersen has moved into the field of open-source professional scientific software development with a high degree of interoperability across the materials science domain, utilizing and developing ontologies. He is currently leading efforts for a CIF (Crystallographic Information File/Framework) ontology alongside tools for implementing and utilizing the Battery Interface Ontology (BattINFO) across interoperability infrastructures and beyond.

Marek Marcinek is a Tenured Professor at Warsaw University of Technology. Born on 22 October 1973 in Bytom. Polish scientist specializing in energy storage systems (Li-ion, Na-ion ...) and entrepreneurship/management education. Marek Marcinek is an alumnus of the Faculty of Chemistry at Warsaw University of Technology, Warsaw School of Economics (Marketing and Management) and Law Department of Warsaw University/University of Poitiers (European Law). He worked in USA (Lawrence Berkeley National Laboratory, CUNY New York), Canada, France and Germany. Marcinek was involved in many worldwide scientific projects. He is a member of the European Research Institute Alistore.

Anna Szczęsna-Chrzan MSc, PhD Candidate is a scientific assistant at the Warsaw University of Technology. Her scientific field covers Li-ion and Na-ion batteries, especially cathodic active materials, electrodes fabrication, and cell balancing. She is also involved in Battery Interface Ontology and batteries production lines engineering. Szczęsna-Chrzan is a certified expert in the project management field. Despite the initial phase of a research career, Anna Szczęsna-Chrzan is involved in many scientific projects, which led to the patent registrations and authorship of scientific papers. She is a member of the BIG-MAP project, part of the Battery 2030+ European Initiative.

Miran Gaberšček is Head of Department for Materials Chemistry at the National Institute of Chemistry, Ljubljana, Slovenia. Since 2009 he has also been Director of Centre of Excellence for Low Carbon Technologies, Ljubljana, Slovenia. Since 2003 he has been affiliated to the University of Ljubljana where he currently holds the position of Full Professor for Materials Science. The primary research field of M. Gaberšček has been novel electrochemical systems for energy storage and conversion (batteries and fuel cells). In particular, he has been focused on explanation of complex charge transport and reaction mechanisms in these devices.

M. Rosa Palacín is a scientist at the Institute of Materials Science of Barcelona (ICMAB-CSIC). Her career has been fully focused on solid-state chemistry and electrochemistry applied to batteries covering a wide diversity of technologies from commercial (e.g., Ni or Li-ion) or precommercial (Na-ion) concepts to new emerging chemistries (Mg, Ca). She has led diverse national and European projects, both public and private funding, and served as co-scientific director of the ALISTORE ERI between 2010 and 2017. She was elected to the Governing Board of Batteries Europe ETIP (Future and Emerging Technologies) in 2019.

Martin Uhrin studied computational physics at Edinburgh University, where he gained his M.Phys. and went on to complete a PhD at the University College London on the topic of high-throughput structure prediction of materials. As a postdoctoral research at EPFL, Martin created the workflow engine that powers the AiiDA materials informatics platform and was responsible for machine learning research to accelerate materials discovery. After a period at DTU researching metal-air batteries, Martin returned to EPFL to conduct independent research using invariant and equivalent machine learning models to enable the inverse design of materials and molecules.

Jesper Friis is a senior scientist at SINTEF in the group of Materials Physics with an associate position at Norwegian University of Science and Technology (NTNU). An important part of his research has been on modelling of metallic alloys and solar cell silicon from the electronic to microstructure scale. The need for robust coupling of software models at different scales made him a central participant in the work conducted by the European Materials Modelling Council (EMMC) on interoperability and the development of the Elemental Multiperspective Material Ontology (EMMO). He has long experienced with EU projects, both as coordinator and participant, and has been an active contributor to EMMC since 2014 where he serves in the Board of Directors.