From Task-Specific AI to Battery Foundation Models

Citation for the original published paper (version of record):
Zhao, J., Cai, M., Li, Y. et al (2026). From Task-Specific AI to Battery Foundation Models. Journal of the Electrochemical Society, 173(15). http://dx.doi.org/10.1149/1945-7111/ae9406

Journal of The
Electrochemical Society

From Task-Specific AI to Battery Foundation Models

To cite this article: Jingyuan Zhao *et al* 2026 *J. Electrochem. Soc.* **173** 150520

Journal of The Electrochemical Society, 2026 **173** 150520

From Task-Specific AI to Battery Foundation Models

Jingyuan Zhao,[1,z] Misheng Cai,[2] Yuqi Li,[3] Shengyu Tao,[4] Andrew F. Burke,[1] and Stephen Harris[5]

[1] *Institute of Transportation Studies, University of California, Davis, Davis, CA 95616, United States of America*
[2] *School of Automotive and Traffic Engineering, Hubei University of Arts and Science, Xiangyang 441053, People's Republic of China*
[3] *Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305, United States of America*
[4] *Department of Electrical Engineering, Chalmers University of Technology, Gothenburg 41296, Sweden*
[5] *Energy Storage and Distributed Resources Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States of America*

Artificial intelligence (AI) is influencing battery research and life-cycle management, spanning materials discovery, electrode and cell design, manufacturing optimization, diagnostics, safety monitoring, recycling, and certification. These advances are enabled by growing battery datasets and learning methods that extract patterns from complex electrochemical systems. Yet most current models remain task-specific and are trained on fragmented datasets, limiting transferability across chemistries, cell formats, operating conditions, and life-cycle stages. Battery foundation models (BFMs) offer a promising but emerging paradigm for battery intelligence. They may be formulated as large, pre-trained, and adaptable architectures that integrate multimodal battery data with physics-informed priors to represent electrochemical structures, states, and dynamics. Unlike direct analogues of general large language models, BFMs should be understood as physics-aware foundation models designed for scientific reasoning under multiscale physical constraints, heterogeneous and incomplete data, and stringent safety and reliability requirements. By combining self-supervised learning with mechanistic electrochemical models, BFMs may learn more interpretable and transferable representations across materials, formats, and duty cycles. If sufficiently validated, such representations could support materials discovery, manufacturing, diagnostics, recycling, and certification. This Review discusses challenges in data fragmentation, validation, interpretability, and computational sustainability, and outlines a roadmap toward responsible BFMs and integrated battery-domain ecosystems across the lifecycle.

[DOI: 10.1149/1945-7111/ae9406]

Manuscript submitted March 27, 2026; revised manuscript received June 29, 2026. Published August 12, 2026.

Abbreviation	Definition
AI	Artificial intelligence
BERT	Bidirectional encoder representations from Transformers
BFM	Battery foundation model
BMS	Battery management system
ECM	Equivalent circuit model
EOL	End-of-life
EU	European Union
EV	Electric vehicle
eVTOL	Electric vertical take-off and landing
FAIR	Findable, accessible, interoperable, and reusable
GNoME	Graph networks for materials exploration
IoT	Internet of things
ISC	Internal short circuit
LFP	Lithium iron phosphate
LLM	Large language model
LoRA	Low-rank adaptation
MAE	Mean absolute error
MAPE	Mean absolute percentage error
NCA	Nickel cobalt aluminum oxide
NCM	Nickel cobalt manganese oxide
non-IID	Not independent and identically distributed
R^2	Coefficient of determination
RMSE	Root mean square error
RUL	Remaining useful life
SEI	Solid electrolyte interphase
SOC	State of charge
SOH	State of health
SSL	Self-supervised learning
TEM	Transmission electron microscopy
X-ray CT	X-ray computed tomography

[z] E-mail: jyzhao@ucdavis.edu

Artificial intelligence (AI) has accelerated progress in battery science and engineering. Over the past decade, task-specific machine learning models have demonstrated practical value in lifetime prediction,[1] electrolyte formulation,[2] fast-charging design,[3] and fault detection.[4,5] Previous reviews of battery AI and studies of battery data landscapes have summarized advances in materials discovery, manufacturing optimization, characterization, state estimation, and safety diagnosis. However, most of these discussions remain organized around specific tasks, life-cycle stages, or modeling pipelines.[6,7] Recent studies further illustrate the deepening integration of batteries and AI, including specialized convolutional Transformer networks for battery health estimation based on transfer learning,[8] resource-efficient convolutional FlashAttention fusion networks for battery capacity estimation,[9] and unsupervised fault diagnosis and localization methods that are compatible with multiple grouping strategies for electric vehicle (EV) battery packs under realistic conditions.[10] Safety-critical studies, including battery fault diagnosis from laboratory to real-world conditions,[11] thermal-runaway characterization,[12] and internal short circuit (ISC) failure analysis,[13,14] also highlight the need for modeling frameworks that can link electrochemical degradation, thermal evolution, fault detection and early warning, and physical safety constraints.

Despite these advances, current battery AI methods remain fragmented. Most models are trained for specific targets, chemistries, cell formats, data regimes, or operating contexts, and their learned representations are often limited in their transferability across materials, cells, packs, and life-cycle stages. The foundation-model paradigm, exemplified by large language models (LLMs) and multimodal models, provides a useful conceptual reference through large-scale pretraining, multimodal alignment, zero-shot and few-shot adaptation, and emerging forms of analogical reasoning.[15–18] Yet battery systems differ fundamentally from language domains, as their behavior is governed by electrochemical kinetics, thermal processes, safety limits, and multiscale degradation dynamics. Accordingly, the central challenge is not the direct application of LLMs to batteries, but the translation of the foundation-model paradigm into physics-aware battery modeling frameworks. In this review, we define battery foundation models (BFMs) as pre-trained, transferable, and physics-aware model families designed to learn from heterogeneous battery data across modalities, scales, and life-cycle stages. Unlike conventional task-specific algorithms, BFMs aim to provide reusable representations that can be adapted with limited retraining to applications such as materials discovery, manufacturing quality control, real-time diagnostics, predictive maintenance, second-life assessment, and recycling. The contribution of this Review is to reframe existing battery AI methods as components of a broader foundation-model-oriented framework for life-cycle-wide battery intelligence, rather than to propose a replacement for them. We clarify the definition, architectural requirements, technical boundaries, risks, evaluation protocols, and development roadmap of BFMs, and discuss how fragmented advances in battery AI could be integrated into transferable, physics-aware, and life-cycle-responsible modeling systems.

The remainder of this review is organized as follows: Section Lessons from Large Language Models examines the capabilities and limitations of LLMs and their implications for physics-constrained, safety-critical BFMs. Section Architecture of Battery Foundation Models outlines key design pillars for BFMs, including multimodal integration, physics-informed modeling, and transferability across chemistries, scales, and operating conditions. Section Applications Across the Lifecycle surveys life-cycle applications in materials design, manufacturing, operation, second-life use, recycling, and policy design. Section Roadmap: Toward Lifecycle-Responsible BFMs discusses challenges related to data fragmentation, generalization, interpretability, certification, computational sustainability, and governance. Section Conclusions proposes a staged roadmap for developing responsible BFMs across the battery life cycle, based on shared datasets, hybrid physics-informed AI architectures, and integrated deployment ecosystems.

Lessons from Large Language Models

A foundation-model perspective suggests that battery intelligence could be advanced by learning from heterogeneous data, including electrochemical signals, imaging, and field measurements (Fig. 1). Unlike language data, however, battery systems are governed by thermodynamic, kinetic, and safety constraints and evolve across coupled spatial and temporal scales. BFMs may therefore need multimodal representation, cross-scale reasoning, and explicit physical constraints to support transferable applications in materials discovery, manufacturing quality control, operational diagnostics, and end-of-life (EOL) assessment.

Transferability and multimodality.—LLMs suggest that large-scale pretraining may support the learning of transferable representations and enable adaptation to diverse tasks under limited supervision. Studies on in-context learning,[19] zero-shot chain-of-thought prompting,[17] multimodal projection alignment,[20] and unified multimodal architectures such as Macaw-LLM [21] have illustrated the potential of foundation models for few-shot adaptation, prompt-based reasoning, and cross-modal understanding. For battery science, however, the key implication is not the direct transfer of language models, but the reformulation of the foundation-model paradigm around electrochemical data, physical constraints, and life-cycle contexts.[22]

Battery data are highly heterogeneous and are distributed across laboratory experiments, field operations, and industrial manufacturing settings.[23,24] Such data include electrochemical measurements, including voltage, current, temperature, and impedance,[25,26] spectroscopic and diagnostic information, for example, electrochemical noise [27,28] and in situ/operando spectroscopy,[29,30] imaging data, such as X-ray computed tomography (X-ray CT) [31–33] and electron microscopy images,[34,35] and fleet-level telemetry that captures charging behavior [36–38] and environmental conditions.[39,40] These modalities capture complementary information about materials, interfaces, cells, packs, and usage environments, but they are often analyzed separately.

The potential transferability of BFMs may therefore depend on aligning multimodal battery data into shared representation spaces that preserve physical meaning across chemistries, cell formats, operating conditions, and life-cycle stages. Recent collaborative AI systems that integrate materials data, electrochemical modeling, and manufacturing knowledge provide early examples of such cross-scale integration.[41] Building on this direction, BFMs could potentially connect voltage curves, impedance spectra, imaging features, manufacturing parameters, and field operational histories to support materials screening, process optimization, degradation diagnosis, safety monitoring, and EOL assessment.

Limitations in reasoning and knowledge.—The limitations of LLMs also provide important cautions for battery applications. Studies have shown that LLM performance can degrade under counterfactual task variants, suggesting potential reliance on narrow heuristics rather than robust reasoning.[42] Other analyses distinguish formal linguistic competence from functional understanding,[43] while studies of chain-of-thought prompting indicate that intermediate reasoning can be unfaithful to the underlying decision process, even when final answers are correct.[44] In electrochemical systems, such reasoning failures can be particularly consequential, as reliable predictions need to remain consistent with thermodynamic limits, kinetic constraints, and conservation laws.

LLMs are also vulnerable to hallucinations and static or outdated knowledge. They may generate confident but incorrect responses, which can be particularly problematic when LLMs are extended to multimodal deployment.[45] In battery contexts, poorly grounded outputs could mislead thermal runaway risk assessment or state of health (SOH) estimation.[46,47] At the same time, recent work that integrates the Internet of Batteries with an LLM-based battery health management framework suggests that few-shot fine-tuning and cross-scenario generalization may support efficient monitoring and prediction of EV battery states.[48] Tool-augmented systems can further improve scientific workflows,[49] but battery applications require physical constraints, domain grounding, and continuously updated data. Thus, LLMs should not be treated as generic reasoning engines for batteries without domain-specific adaptation and validation.

Implications for battery foundation models.—The main lesson from LLM research lies in the foundation-model paradigm of pre-training, adaptation, and transfer, rather than in language generation itself. Reusable representations,[50] instruction tuning,[51] supervised fine-tuning,[52] alignment with human feedback, retrieval augmentation,[53] and parameter-efficient methods such as low-rank adaptation (LoRA) [54] provide useful methodological references for BFMs. However, their use in battery science needs to be translated into battery-specific workflows.

For BFMs, model development could begin with battery-specific data acquisition pipelines that systematically integrate textual and non-textual sources, including scientific literature, patents, materials databases, cycling records, voltage and impedance curves, manufacturing logs, battery management system (BMS) data, field operation records, failure reports, microscopy data, spectroscopy data, and simulation outputs. These data should ideally be accompanied by structured metadata, such as chemistry, cell format, temperature, C-rate, state of charge (SOC), SOH, testing protocol, uncertainty, and provenance, which can help models learn electrochemically meaningful relationships rather than purely textual or statistical associations.[23] In this context, instruction tuning could help models follow battery-specific task descriptions and experimental protocols. Alignment with human feedback could improve expert-facing

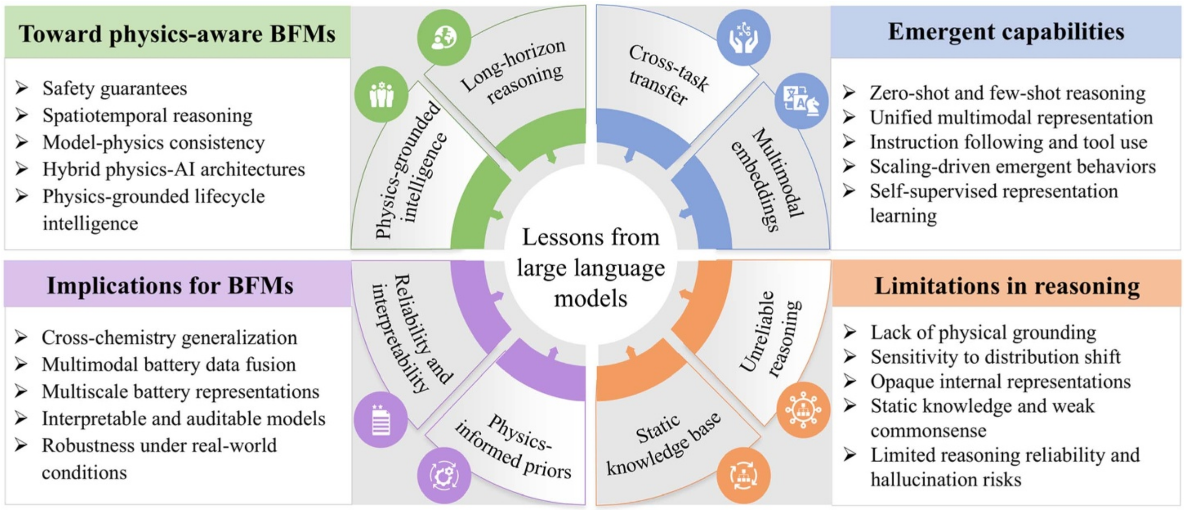

Figure 1. Lessons from large language models for battery foundation models. The emergent capabilities of LLMs, including transferability, zero-shot generalization, and multimodal learning, suggest the value of large-scale pretraining as a conceptual reference for BFMs. However, limitations related to reasoning reliability, static knowledge, and hallucinations motivate the development of physics-aware BFMs that remain responsible across the battery life cycle.

question-answering and decision-support tasks. Recent work on tool-adaptive scientific LLMs further suggests that models can learn when to rely on direct reasoning versus external tools, a capability that may be useful for BFM workflows involving literature retrieval, simulations, and data analysis.[55] Multimodal fine-tuning may also be needed to align language with voltage curves, impedance spectra, microscopy images, manufacturing parameters, and operational time series. Thus, the key lesson from LLMs for battery science is the redesign of foundation-model workflows around battery-specific data acquisition, domain-aware fine-tuning, and physics-constrained adaptation.

Toward physics-aware foundation models.—Battery systems call for models that can represent physical processes governed by thermodynamic, kinetic, and safety constraints, rather than architectures that learn only statistical patterns. Future BFMs may therefore need self-supervised pretraining across multimodal data streams,[56] together with electrochemical governing equations,[57] digital-twin representations,[58] and explicit physical constraints. In addition, recent work on a domain-adaptive battery capacity prediction framework suggests that impedance-based physical constraints can help capture mechanistic information during random discharge processes, supporting nondestructive discharge capacity prediction, cross-scenario generalization, and aging mechanism interpretation.[59] By enforcing physical feasibility, such constraints may reduce the risk of unphysical outputs, spurious correlations, and inconsistent recommendations, thereby supporting foundation-model approaches that are more interpretable, trustworthy, and suitable for battery science and engineering.

Architecture of Battery Foundation Models

The capabilities of BFMs depend not only on data scale, but also on architectures that enable multimodal integration, physical consistency, and transfer across contexts. In this review, a BFM is defined as a pre-trained model or model family designed to learn reusable representations from heterogeneous battery data across modalities, scales, and life-cycle stages, and to support adaptation to multiple battery tasks with limited task-specific retraining. Operationally, such models can be characterized by three criteria: (i) multimodal and cross-scale data integration beyond a single task-specific dataset; (ii) physics-informed and safety-constrained representation learning that incorporates electrochemical, thermal, or operational constraints; and (iii) scalable transfer across chemistries, cell formats, system levels, and operating contexts. This definition helps distinguish BFMs from existing battery modeling approaches. Physics-informed machine learning, digital twins, multimodal learning, multitask or transfer-learning models, and integrated task-specific workflows each provide important components of a future BFM ecosystem. However, they should not be regarded as equivalent to BFMs unless these capabilities are integrated through shared pre-trained representations that support broad adaptation across tasks and life-cycle stages. Thus, BFMs should be viewed not as replacements for existing approaches, but as an integrative and still largely aspirational framework that incorporates and extends them. Table I clarifies this operational boundary, and the following subsections discuss three design pillars: multimodal integration, physics-informed modeling, and scalability and transferability (Fig. 2).

Multimodal integration.—The potential value of BFMs may lie in their ability to integrate heterogeneous battery data across modalities, scales, and life-cycle stages. Relevant data sources include laboratory cycling tests, in-service telemetry,[60] finite-element simulations, X-ray CT,[33] in situ transmission electron microscopy (TEM), and in situ/operando spectroscopy.[61] Looking beyond lithium-ion technologies, in situ and operando characterization can reveal electrochemical reactions, materials evolution, interfacial degradation, and multiscale dynamics in next-generation battery systems. These methods provide a critical experimental basis for improving the performance and safety of solid-state and post-lithium-ion batteries and for guiding their future development.[62–64] However, such heterogeneity makes multimodal integration a methodological challenge rather than simple data aggregation: transferable representations for BFMs may depend on the alignment, fusion, and validation of heterogeneous data streams.

In practice, multimodal integration involves three steps. First, data should be aligned across temporal, spatial, scale-related, and semantic dimensions.[65] Electrochemical time series may be synchronized by cycle number, SOC, SOH, aging stage, or testing protocol, whereas imaging and spectroscopic data should be linked to electrode regions, degradation modes, or observation windows. Field data further require metadata harmonization, including chemistry, cell format, temperature, charging protocol, duty cycle, and environmental conditions. Second, fusion should occur at appropriate representational levels. Existing taxonomies of multimodal learning classify fusion strategies at the data, feature, and output levels, and

Table I. Operational distinction between BFMs and related battery modeling approaches.

Approach	Representative battery examples	Data scope	Transfer/adaptation	Physical or safety constraints	Current status
Task-specific battery models	SOC estimation, SOH/RUL prediction, fault detection	Single task, dataset, chemistry or operating regime	Limited; usually requires retraining	Optional or task-dependent	Mature application basis, but lacks general-purpose representations
Physics-informed machine learning	SOH prediction with electrochemical constraints; thermal-risk modeling	Usually task-or system-specific datasets	Limited to the modeled task or domain	Explicitly embedded in loss functions, architectures, or post-processing	Provides mechanisms for physical constraint integration
Digital twins	Cell-, pack-, or manufacturing-process simulation, monitoring and control	Dynamic data from selected systems or processes	Reuse through calibrated models or module outputs	Usually includes physical, thermal or operational constraints	Provides dynamic system-level representations
Multimodal learning models	Fusion of voltage, impedance, imaging, spectroscopy or operational data	Multimodal data, usually for a specific task	Depends on modality alignment and domain shift	Usually limited unless added explicitly	Provides heterogeneous data-fusion capability
Multitask or transfer-learning models	Cross-dataset SOH prediction, domain-adapted degradation modeling	Related tasks, datasets or domains	Moderate, but limited by domain shift	Usually limited or added post hoc	Provides cross-task or cross-domain adaptation capability
Integrated task-specific workflows	Multiscale modeling or battery digital-twin workflows linking lifecycle modules	Selected lifecycle stages connected by workflow logic	Reuse through module outputs, not necessarily shared pre-trained representations	Present in selected modules	Provides modular lifecycle linkage
Battery foundation models	Prospective lifecycle model trained on lab, manufacturing, field, and EOL data	Heterogeneous data across modalities, scales, chemistries and lifecycle stages	Broad adaptation to multiple tasks with limited retraining	Explicit electrochemical, thermal, operational or safety constraints	Integrates the above capabilities; remains largely aspirational

Core architectural pillars of battery foundation models

Multi-modal integration	Physics-informed modeling	Scalability and transferability
Electrochemical signals	**Physics-embedded models**	**Cross-scale transfer**
▪ High-rate charging signatures	▪ Thermal–mechanical coupling	▪ Cell to module feature aggregation
▪ Electrochemical time-series profiles	▪ Electrochemical and aging PDEs	▪ Module to pack state mapping
▪ Polarization and relaxation dynamics	▪ Safety envelopes and protection logic	▪ Pack to fleet-level risk modeling
▪ Impedance and frequency-domain features	▪ Reduced-order mechanistic surrogates	▪ Stress propagation across scales
Imaging and spectroscopy	**Hybrid learning loops**	**Cross-chemistry generalization**
▪ SEM/TEM nanoscale morphology	▪ Digital-twin–driven updates	▪ Na-ion and other emerging chemistries
▪ X-ray tomography microstructure	▪ Active experiment–model iteration	▪ Graphite, Si-rich and alloy anodes
▪ Operando diffraction and spectroscopy	▪ Physics-informed neural networks	▪ Li-metal and solid-state architectures
▪ Pore-scale and binder-network mapping	▪ Physics-regularized learning	▪ LFP, NCM/NCA and high-Ni cathodes
Field and contextual data	**Constraints and interpretability**	**Cross-application adaptation**
▪ Fleet telemetry and load patterns	▪ Mechanistic interpretability	▪ Grid-scale and distributed storage
▪ Charging and grid-interaction data	▪ Conservation-law constraints	▪ Extreme climates and high-rate regimes
▪ Climatic and geographic variability	▪ Causal degradation reasoning	▪ Light- and heavy-duty vehicle platforms
▪ Operational events and maintenance logs	▪ Certifiable and safe model behavior	▪ Aviation, marine and off-highway sectors

Figure 2. Core architectural pillars and key design considerations for battery foundation models. BFMs may depend on three mutually reinforcing elements: (i) multimodal integration of electrochemical, imaging, and field data; (ii) physics-informed modeling through embedded equations, hybrid learning loops, and interpretability constraints; and (iii) scalability and transferability across system levels, chemistries, and application contexts.

highlight their roles in cross-modal understanding, robust fusion, and generalization.[66] Third, multimodal integration should be evaluated beyond task-level accuracy. Modality ablation, missing-modality tests, cross-dataset validation, and laboratory-to-field transfer tests are needed to assess whether models use complementary information and generalize across battery contexts.[60,67,68] Through these procedures, multimodal integration could become a key architectural requirement for BFMs.

Physics-informed modeling.—Although multimodal integration expands the data available to BFMs, its value remains limited without physical consistency. Battery behavior is governed by electrochemical kinetics, charge conservation, mass transport, and thermal processes. Trustworthy prediction, especially in safety-critical scenarios, therefore requires physical principles to be embedded within model architectures, training objectives, or output constraints. Physics-informed neural networks and related hybrid models indicate that incorporating empirical degradation trends and state-space constraints,[69] together with laboratory measurements and field data,[70] can improve SOH prediction across battery chemistries and operating conditions. Similar physics-informed approaches can also improve inverse thermal modeling from sparse observations,[71] illustrating their value for robust and interpretable battery modeling.

At the electrochemical level, BFMs should be designed to capture relationships associated with ion transport, reaction kinetics, and aging mechanisms. Mechanistic descriptions, such as nonlinear electrochemical impedance theory [72] and coupled electro-thermal-aging models,[73] provide important physical anchors for representation learning. Beyond predictive accuracy, safety constraints, such as voltage windows,[74] internal pressure thresholds,[75] and thermal stability margins,[76] may be explicitly incorporated to reduce the risk of physically infeasible or unsafe recommendations. In addition, hybrid physics-AI training loops provide a practical pathway for embedding these constraints. By penalizing deviations from governing equations [77] or using digital-twin environments as physics-consistent testbeds,[78] such approaches may reduce reliance on labeled data, improve interpretability, and limit spurious correlations. For BFMs, the goal is not merely to improve prediction accuracy, but to combine data-driven adaptability with mechanistic consistency, thereby supporting more trustworthy model outputs across the battery life cycle.

Scalability and transferability.—For BFMs, scalability should refer less to a universal model covering all battery phenomena than to transferable representations across clearly defined system levels and task boundaries. At the cell level, relevant factors include chemistry, form factor, electrode design, electrochemical response, and degradation. At the pack level, transfer should account for electrical interconnection, thermal management, sensing, balancing, and BMS control, because system-level integration in electric mobility couples electrochemical, structural, thermal, electrical, and management functions rather than simply scaling cell-level behavior.[79] At the fleet or deployed-system level, models would need to consider real-world usage, climate, charging behavior, maintenance, and operating conditions. Fragmented data ecosystems, inconsistent standards, and limited sharing mechanisms remain major obstacles to such cross-scale modeling.[23]

Transferability is further constrained by chemistry, design, and deployment context. Lithium iron phosphate (LFP), nickel cobalt manganese oxide (NCM) and nickel cobalt aluminum oxide (NCA), lithium-metal, sodium-ion, and solid-state batteries differ in material composition, voltage behavior, safety characteristics, and aging pathways. Cell format and electrode design also influence heat generation, degradation, and capacity fade.[50] Laboratory data are usually collected under controlled conditions, whereas field data reflect uncontrolled usage, lower sensor accuracy, incomplete logging, and sparse ground-truth validation.[80] Recent transfer-learning work has

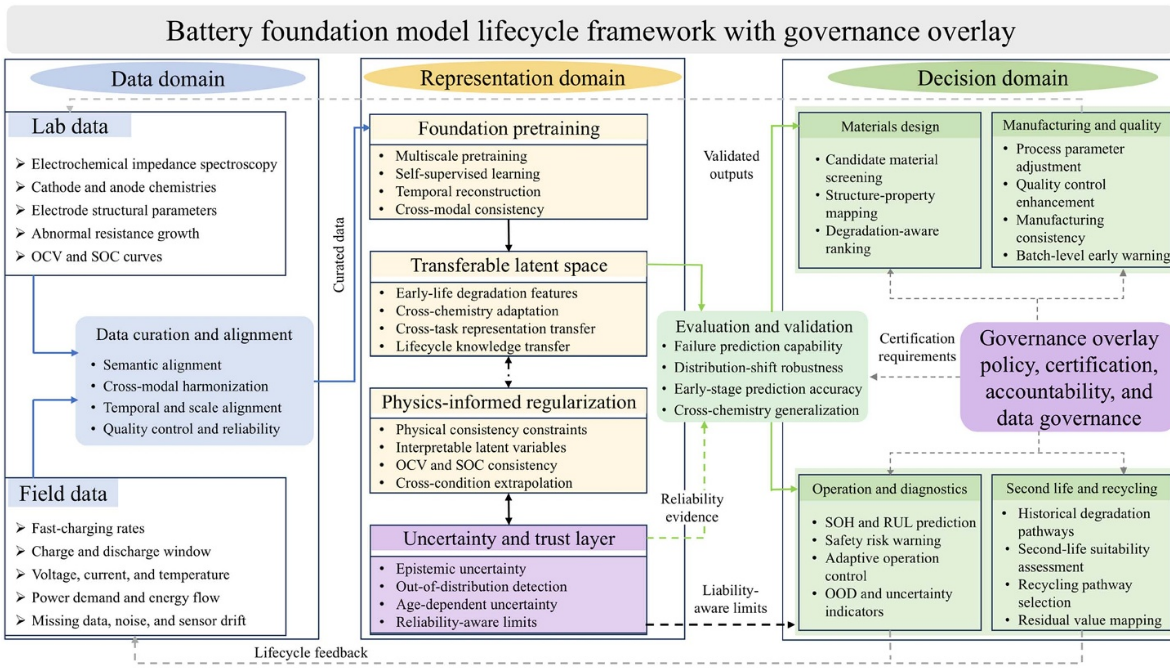

Figure 3. Lifecycle framework for battery foundation models with governance overlay. Lab and field data are curated and aligned to support foundation pretraining, transferable representation learning, physics-informed regularization, uncertainty estimation, and validation. The validated outputs could support lifecycle applications in materials design, manufacturing quality control, operation and diagnostics, second-life assessment, and recycling. Governance constraints related to policy, certification, accountability, and data governance guide the validation and deployment process, while lifecycle feedback supports continual model updating.

reported SOH estimation across LFP, NMC, and NCA chemistries and from cells to packs under controlled validation conditions.[81] Robust BFMs nevertheless require validation under explicit domain shifts.

Scalability is also limited by data availability, proprietary constraints and computational cost. Public battery datasets differ in protocols, variables and reporting formats,[82] while detailed geometry, manufacturing parameters and material formulations are often commercially sensitive.[83] Some safety events, such as thermal runaway triggered by thermal-management failures, require system-level models beyond chemistry or degradation data alone.[7] In addition, large-scale pretraining demands substantial memory, computation and parallel training infrastructure. Thus, BFMs may support cross-scale battery analysis when data are well organized, system boundaries are explicit, task transfer is validated, and computational strategies remain practical.

Applications Across the Lifecycle

BFMs are a prospective framework for connecting fragmented modeling tasks across the battery life cycle. Existing studies have advanced task-specific applications in materials discovery, manufacturing quality assurance, operational diagnostics, second-life assessment, and recycling decision support. Building on these developments, BFMs may provide shared representations that link laboratory measurements, manufacturing data, operational telemetry, and EOL information. If rigorously validated, they could support coordinated decision-making across materials design, manufacturing control, operational diagnostics, second-life deployment, and recycling. With auditable, traceable, and interpretable outputs, BFMs may also contribute to certification, policy communication, and life-cycle governance. The full capabilities of BFMs across heterogeneous datasets, chemistries, scales, and deployment contexts remain to be demonstrated (Fig. 3).

Materials discovery and design.—Materials discovery is one of the near-term prospective applications of BFMs. Existing task-specific studies have improved screening efficiency through high-throughput quantum chemistry workflows,[84] data-driven catalyst discovery for lithium-sulfur batteries,[85,86] cycle-life prediction from capacity-voltage profiles,[87] and nonlinear modeling of electrochemical stability windows.[88] However, these approaches remain largely confined to specific materials, tasks or operating regimes, with limited transferability across chemistries and scenarios. Against this background, BFMs may provide a route to connect these fragmented advances, but this capability should be viewed as prospective rather than established. Such representations, constructed by integrating atomistic simulations,[89] spectroscopic measurements,[90] cryogenic electron microscopy characterization,[91] and electrochemical performance testing could, in principle, support faster screening, property prediction, and constrained inverse design. Generative models already show promise for proposing materials with target functionalities,[92,93] but BFM-based inverse design would require physical and chemical constraints, such as thermodynamic feasibility [94] and mechanical-chemical degradation pathways,[95] to reduce the risk of implausible candidates. A unified BFM that robustly combines multimodal materials representation, property prediction and constrained inverse design across battery chemistries remains to be demonstrated.

Manufacturing and quality assurance.—Manufacturing quality assurance determines whether laboratory advances can be translated into scalable battery products. Although gigafactory expansion has increased the need for resilient and efficient production,[96] many quality control systems remain rule-based and rely on post hoc inspection. Recent studies show that richer process signals, including calendaring signatures,[97] ultrasonic leakage detection,[98] impedance-based analysis,[99] and automatic optical inspection,[100] can improve defect detection and process monitoring. For BFMs, manufacturing may offer an opportunity to integrate heterogeneous process signals into shared representations for defect detection, process monitoring, and quality prediction. Continuous in-line monitoring and closed-loop feedback, such as slurry viscosity control, may reduce downtime, scrap rates, and batch variability.[101] In addition, the economic and environmental stakes are substantial, as manufacturing defects and scrap can increase material demand, production cost, and life-cycle impacts.[102] However, BFM-enabled manufacturing control would require representative plant-scale datasets, real-time validation, and evidence that model recommendations remain reliable under changing materials, equipment settings, and production conditions.

Operation and diagnostics.—In deployed systems, batteries operate under stochastic and heterogeneous conditions that differ substantially from laboratory testing.[103,104] One study indicates that EV battery lifetime can vary markedly by climate, with lifetimes approximately 30% longer in cold regions than in hot regions,[105] and commercial aging data show nonlinear dependence on operating conditions.[106] These findings motivate models that combine operational signals with mechanistic degradation knowledge under real-world uncertainty. Meanwhile, state estimation remains central to operation and diagnostics. Conventional equivalent circuit models (ECM),[11,107] Kalman-filter-based estimators,[108] and regression methods [109] can perform well in controlled settings, but may lose robustness under fleet-level variability and unseen conditions. Deep transfer-learning architectures that combine convolutional feature extraction with self-attention have achieved cross-chemistry SOH diagnostics at millisecond-level latency and demonstrated robustness using noisy fleet-operational data.[110] Multimodal frameworks suggest that shared representations could improve SOC, SOH, and RUL estimation when multiple data streams are available.[60] Task-specific work using real-world operational data from 300 EVs further shows that dynamic graph learning and interpretable temporal decomposition can support SOH estimation under cell-level heterogeneity and user-specific charging variability.[111] Future BFMs may also support anomaly detection and fault diagnosis by linking subtle precursors, such as microcracking,[112,113] local lithium plating,[114] or abnormal electrolyte consumption,[115] with calibrated uncertainty and low false-alarm rates. These signals may contain useful diagnostic information, but it has not yet been demonstrated that a unified BFM can reliably detect such precursors in real time across different batteries and operating conditions.

For safety monitoring, events such as thermal runaway [12,46,116] and ISCs [117] may threaten individual cells and, under unfavorable conditions, propagate into system-level failures. Recent work has introduced bidirectional encoder representations from transformers (BERT)-style self-supervised pretraining frameworks that tokenize time series and use pointwise masking to learn distributionally robust, context-aware temporal representations, with further improvements from battery metadata.[14] This provides task-specific evidence that self-supervised representation learning may improve fault detection from battery time-series data. However, such studies should be interpreted as task-specific evidence rather than demonstrations of a fully unified BFM for battery safety assurance. These studies instead suggest a possible direction in which future BFMs may integrate fault detection, uncertainty estimation, pack-level context, and operational constraints to support earlier warning and more adaptive safety management. Capabilities such as adaptive derating or dynamic pack reconfiguration remain prospective and would require validation under representative fault scenarios, sensor noise, pack heterogeneity, and safety-critical deployment conditions.

Second life and recycling.—Second-life utilization and recycling are essential for closing battery material loops and reducing life-cycle impacts. Existing studies have advanced the valorization and recovery of retired batteries from multiple perspectives, including economic assessment, pathway optimization, state identification, and rapid diagnostics. For example, degradation-aware economic models can improve second-life allocation,[118] multi-objective optimization can balance reuse and material recovery,[119] and diagnostic methods such as generative-learning-assisted SOH estimation,[120] privacy-preserving cathode-material identification,[121] and pulse-test-based health diagnosis [122] can support fast screening, graded utilization, and efficient recycling of EOL batteries. Nevertheless, these methods remain largely stage-specific and do not yet provide a unified framework linking retired-battery diagnosis, second-life triage, disassembly risk assessment, and recycling-route selection.

For the circular battery economy, future BFMs may provide a basis for learning shared health representations from in-service telemetry, laboratory diagnostics, EOL inspection, and recovery data, provided that they are developed and validated using sufficiently diverse life-cycle data. Such representations may help connect retired-battery diagnosis to second-life triage by informing the identification of cells and packs that are potentially suitable for reuse, repair, direct recycling, or further inspection. They could also provide inputs for disassembly risk assessment by capturing degradation state, structural heterogeneity, and safety-relevant signatures, although robotic and AI-assisted disassembly remains challenging for retired packs with diverse designs and safety risks.[123,124] BFMs could also be explored for recycling-route selection by linking aging signatures and post-disassembly state assessment to realized material yields under alternative recovery pathways, such as direct cathode regeneration,[125] hydrometallurgical leaching [126] and pyrometallurgical processing.[127] In this way, BFMs may eventually contribute to decision-support workflows across the chain from diagnosis to triage, disassembly, and recycling; however, their practical value for building more efficient, safe, and sustainable battery ecosystems remains to be demonstrated under heterogeneous chemistries, uncertain usage histories, inconsistent EOL inspection protocols, and diverse recycling pathways.

Policy, certification, and accountability.—Large-scale battery deployment creates institutional demands for transparency, compliance and accountability. Digital battery passports mandated by the European Union (EU) Battery Regulation [128] and extended producer responsibility policies are reshaping requirements for battery traceability and lifecycle management.[129] Insurers and financial institutions likewise require robust methods to quantify operational risk and residual asset value. Meeting these requirements requires more than performance prediction alone. It calls for intelligent systems that are auditable, verifiable, and capable of providing life-cycle-spanning insights. BFMs may become one component of this governance architecture. By integrating heterogeneous data streams, including laboratory diagnostics,[116] fleet-level operational telemetry,[38,130] and EOL processing outcomes,[131] BFMs could support analytics capabilities for battery passports, enabling transparent records of chemical composition, performance histories, and degradation trajectories. Such records could help address regulatory and compliance needs while also supporting new business models, such as battery-as-a-service, whose value propositions depend on traceable and verifiable asset performance. In the battery domain, uncertainty-aware BFMs could support battery-passport analytics, certification, residual-value assessment, and risk-differentiated decision-making, provided that their outputs remain interpretable, traceable, and independently auditable.

Challenges.—The development of BFMs faces challenges in data availability, model reliability, computational cost, and governance. Battery datasets remain fragmented, and industrial confidentiality limits access to representative manufacturing, field operation, and failure data. Models that rely mainly on statistical patterns without sufficient physical constraints may fail under distribution shifts, especially in safety-critical applications. Limited transparency further complicates verification, traceability, and deployment. In addition, large-scale pretraining raises concerns about computational cost, energy use, and carbon footprint. Addressing these risks

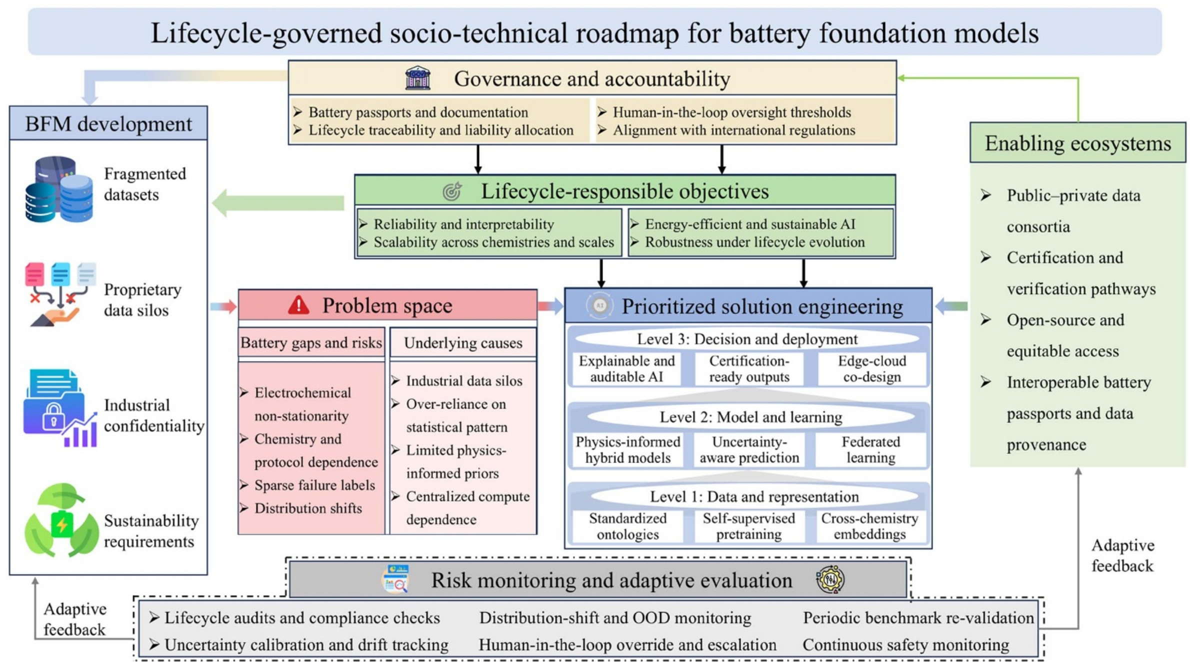

Figure 4. Lifecycle-governed socio-technical roadmap for battery foundation models. The roadmap illustrates how BFM development constraints, including fragmented datasets, proprietary data silos, industrial confidentiality, and sustainability requirements, may give rise to technical and governance risks. These risks could be addressed through prioritized solution engineering across data representation, model learning, and decision deployment, under lifecycle-responsible objectives and governance accountability. Risk monitoring, adaptive evaluation, and enabling ecosystems could provide feedback for continuous model updating, certification, interoperability, and lifecycle-responsible deployment.

may require physics-informed modeling, standardized evaluation protocols, transparent data and model pipelines, and collaborative mechanisms for accessing high-quality battery datasets (Fig. 4).

Data fragmentation and accessibility.—For BFMs, the data challenge concerns not only scale and access, but also physical incompleteness and life-cycle fragmentation. Battery data arise from materials development, laboratory testing, manufacturing, pack integration, and field operation, but they are often collected by different stakeholders using different protocols, sampling rates, quality control criteria, and reporting formats.[23,82] Laboratory measurements can reveal electrochemical, structural, and failure mechanisms, for example through synchrotron X-ray imaging and thermal imaging of thermal runaway,[132] but they are typically limited in scale and operating coverage. Manufacturing data capture electrode processing, coating uniformity, drying, calendering, and microstructural defects that affect electrode quality, cell consistency, and early-life defects. However, these data are often treated as core industrial assets.[133] Field data are essential for aging analysis but often suffer from uncontrolled conditions, limited sensing, incomplete logging, and infrequent validation.[80] Thus, battery data remain fragmented across scales, modalities, and causal pathways.

This fragmentation may weaken representation learning. Similar external signatures, such as capacity fade, power fade, and impedance growth, may arise from different mechanisms, including solid electrolyte interphase (SEI) growth, lithium plating, loss of active material, particle cracking, or electrolyte depletion.[12] BFMs trained mainly on accessible terminal signals may therefore learn surface correlations rather than degradation mechanisms. Key safety events, such as thermal runaway, are also rare, nonlinear, and costly to collect. Data scarcity can make deep learning models prone to overfitting and may lead to inadequate representations of safety risk, limiting their ability to model extreme failure boundaries.[134] Industrial confidentiality further restricts the coverage of BFM pretraining corpora. Manufacturing and quality control data often involve material formulations, process windows, equipment parameters, yield control, and supply chain information, and they are rarely public. Federated and privacy-preserving learning may enable collaborative modeling without centralizing raw data,[135] but they do not remove the technical barriers. Data across manufacturers, plants, vehicle platforms, and chemistries are likely to be highly non-IID, a setting in which federated models often underperform centralized models and show reduced stability and generalization.[136] Federated training may also introduce communication overhead for large multimodal BFMs because of repeated parameter or gradient exchange.[137] Moreover, gradients are not inherently secure and can leak private training samples.[138] Thus, a central challenge is to develop physically meaningful and transferable BFMs when data cannot be fully shared, key internal states are only partially observable, manufacturing variables are missing, and cross-institutional data remain inconsistent.

Failure modes and robustness safeguards.—BFMs should not be regarded as inherently robust or unbiased. They may be trained on heterogeneous data streams that differ in modality, scale, protocol, noise level, and operating-condition distributions. Without explicit provenance, metadata, and causal factors, models may capture superficial associations rather than electrochemically meaningful relationships. Physical constraints may improve interpretability and extrapolative reliability.[139] However, when laboratory and field data diverge because of differences in aging pathways, thermal management conditions, load profiles, or sensor errors, overly rigid physical constraints may suppress rare but valid observations, whereas overly flexible models may fit noise or biased empirical trends. BFMs may therefore require soft constraints, source-aware learning, conflict detection, and uncertainty quantification, rather than simple hard-coding of physical laws across all data regimes.

Training-data bias is another critical risk. Even integrated datasets remain limited for solid-state batteries, emerging electrolytes, low-temperature operation, abuse conditions, long-term degradation, and real-world failures.[140] Moreover, data from standard operating conditions are typically more abundant than data from extreme temperatures, fast charging, high loading, or practical use environments. Without appropriate safeguards, BFMs may perform well for common chemistries and standard conditions while producing overconfident or misleading predictions for low-resource chemistries, abnormal operating regimes, or safety-critical scenarios. Although fully established BFMs for batteries are not yet available, recent examples of AI-driven materials discovery and battery manufacturing provide quantitative evidence for BFM-relevant capabilities and underscore the need for systematic scrutiny. For example, graph networks for materials exploration (GNoME) demonstrated the scalability of AI-driven crystal discovery by predicting 2.2 million new crystal structures and identifying 380 000 stable candidate materials.[141] Similarly, the integration of AI models with cloud-based high-performance computing enabled the screening of more than 32 million candidate materials within 80 h, predicting approximately 500 000 potentially stable materials and identifying 18 candidate compositions for battery solid electrolytes, with the leading candidate experimentally validated.[142] Beyond materials screening, AI-guided iterative optimization has also been applied to graphite-based anode development using the Citrine Platform. Starting from noisy and incomplete industrial data, this workflow improved manufacturing reliability to 100% successful cell production, increased the fraction of cells delivering at least 350 mAh g^{-1} from 28.4% to 84.8%, and raised capacity retention from 42.1% to 97.3%.[143] These examples suggest that AI models are beginning to move from academic prediction tools toward practical materials discovery and battery-development workflows. They also show why data provenance, training distributions, physical consistency, uncertainty, and external validation should be examined carefully when such methods are extended to high-throughput screening or closed-loop optimization.

Generalization across chemistries and contexts.—A central technical challenge for BFMs lies less in performing well on large pooled datasets than in generalizing across chemistries, cell formats, and deployment contexts while preserving the mechanisms that make each domain distinct. Models trained on one chemistry may fail when applied to another because voltage profiles, degradation pathways, safety margins, and operating windows are chemistry-dependent.[144] Similar voltage, resistance, or capacity signatures may correspond to different internal processes in LFP, high-nickel NCM and NCA, lithium-metal, or solid-state batteries, creating the risk of false generalization.

Generalization is further complicated by cell format and system context. Cylindrical, prismatic, and pouch cells differ significantly in geometry, packaging format, heat dissipation pathways, and mechanical stability.[79] A diagnostic feature calibrated on coin cell or single-cell experiments may therefore become unreliable at the module or pack level, where thermal gradients, cell-to-cell variability, balancing strategies, and BMS control alter the observed response. Deployment context adds another layer of shift. Road EVs, stationary storage, and electric vertical take-off and landing (eVTOL) platforms impose different power, energy, safety, and lifetime requirements. Transportation-focused battery integration studies further show that architectural compactness, safety resilience, serviceability, recyclability, and regulatory accountability are application-dependent trade-offs.[145] For instance, eVTOL batteries require high-power burst capability during takeoff and vertical flight under stringent mass and safety constraints.[146] Generalization should therefore be evaluated not only by average prediction error across datasets, but also under context-specific physical and operational constraints.

For BFMs, the key question is how to separate invariant electrochemical structure from context-dependent variation. Representations that are too general may obscure chemistry-specific degradation pathways, whereas representations that are too specialized may fail to transfer. This tension can lead to negative transfer, mis-calibrated uncertainty, and overconfident predictions in new regimes. Physics-informed learning, transfer learning, and domain adaptation may help constrain representations across materials, geometries, and boundary conditions,[147] but they do not remove the need for careful validation under explicit domain shifts. Therefore, the challenge lies not only in reducing the data required for new chemistries or applications, but also in defining when transfer is valid, when adaptation is required, and when a BFM should flag or abstain from extrapolation beyond its learned physical and operational support.

Benchmarking, evaluation metrics, and validation protocols.—Because BFMs remain an emerging paradigm, their evaluation should not be reduced to predictive accuracy on a single battery dataset or task. At present, no widely adopted benchmark has yet been established for BFMs. Future evaluation could therefore build on existing battery datasets, data format initiatives, and battery lifetime prediction benchmarks, while gradually extending toward multimodal, transferable, and physics-aware assessment.[23,82,140,148] A practical evaluation framework could include three layers. First, benchmark tasks should cover not only SOH, RUL, and capacity fade prediction, but also anomaly detection, safety early warning, degradation mode identification, manufacturing quality assessment, and second-life decision support. Second, predictive metrics such as mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2) should be combined with safety-oriented and reliability-oriented metrics, including false-negative rate, warning lead time, uncertainty calibration, and confidence interval coverage. Third, validation protocols should move beyond random train-test splits, for example by holding out a specific data source, chemistry, cell format, or testing protocol during training and using these unseen conditions to evaluate cross-domain generalization. In addition, modality ablation and missing-modality tests can assess whether models effectively use complementary multimodal information;[60,67] physics-consistency checks can examine whether learned representations and predictions conform to degradation mechanisms and known electrochemical constraints;[149] and laboratory-to-field transfer tests can evaluate applicability under real-world operating conditions.[68]

Interpretability, trust, and certification.—For BFMs to move from research prototypes to deployable tools, accuracy alone is insufficient. In safety-critical applications, predictions need to be interpretable, physically consistent, and auditable. This is especially important in high-risk systems such as electric aviation, where overheating, thermal runaway, or power loss can have severe consequences.[150] Even if a BFM produces accurate outputs, deployment would remain difficult if engineers, operators, insurers, or regulators cannot verify the basis of its decisions. The interpretability challenge is distinctly battery-specific. For black-box decision-support systems, prior work has shown that interpretability helps users understand model outputs, select appropriate methods, and scrutinize model decisions.[151] BFMs may therefore need to provide not only predictions of SOH, RUL, or safety risk, but also indications of whether these predictions are consistent with plausible electrochemical, thermal, mechanical, or system-level mechanisms. Similar observable signals may arise from lithium plating, loss of active material, impedance growth, thermal imbalance, or BMS effects, and failure to distinguish these mechanisms could lead to misleading interventions.

Trustworthy battery AI depends on balanced optimization of performance, safety, cost, privacy, interpretability, and deployable governance.[152] Without explicit physical constraints, priority rules, or safety boundaries, BFMs may converge toward recommendations that are statistically effective but operationally unsafe. Certification further raises the threshold for deployment. Regulatory frameworks such as DO-311A for aviation,[153] ISO 26262 for automotive functional safety,[154] and the EU battery regulation [128] require reproducibility, traceability and evidence for safety-relevant decisions. For BFMs, this means that outputs may need to include uncertainty bounds, provenance records, validation evidence, and auditable decision rationales. When confidence is low, inputs fall outside the training domain, or physical constraints are violated, the system could be designed to switch to conservative policies, trigger human review, or fall back to rule-based or mechanistic models.

Constrained convergence and sustainable scaling.—Sustainable scaling of BFMs should be assessed not only in terms of computational efficiency, but also in terms of their environmental and social implications.[155,156] Battery decisions involve competing priorities, including fast charging, cycle life, power capability, thermal safety, cost, and recyclability. Fast-charging studies, for example, show that charging time, cycle life, and voltage constraints form a typical multi-objective optimization problem.[3] A model optimized only for average prediction accuracy may converge toward statistically efficient but physically implausible or unsafe representations. This risk may increase in multitask BFMs, where materials discovery, manufacturing control, SOH estimation, safety early warning, and recycling assessment may impose conflicting objectives. Physics-informed losses, safety-aware objectives, uncertainty penalties, constraint satisfaction, Pareto front analysis, and task priority rules may therefore be needed.

Computational scaling also raises sustainability concerns. Scaling studies of large neural network training provide useful risk bounds. Training emissions depend strongly on model size, hardware efficiency, data center efficiency, and electricity carbon intensity.[157] The BLOOM analysis shows how sensitive estimates are to system boundaries: for a 176-billion-parameter model, final training emissions were estimated at 24.7 t CO_2 when only dynamic power was counted, but at 50.5 t CO_2 when broader operational and equipment-related processes were included.[158] Inference at scale may also contribute substantially to emissions.[159] Sustainability assessment would therefore need to include training, inference, retraining frequency, and hardware life-cycle costs. The environmental justification for BFMs would need to rest on evidence that downstream benefits, such as reducing scrap rates, extending battery lifetime, improving safety, or enhancing reuse and recycling, outweigh computational and hardware footprints.

Deployment constraints are equally important. Many BFM applications may require low-latency, resource-constrained inference in gigafactory control systems, fleet platforms, or onboard BMS hardware. BMS applications, in particular, require real-time reliability, limited memory use, and implementable algorithms for state estimation and safe operation.[160] Thus, computational sustainability could be evaluated not only by training energy, but also by life-cycle deployment cost, edge inference feasibility, and the ability to maintain calibrated uncertainty under compression or adaptation. In addition, compression methods such as pruning, distillation, LoRA, and modular architectures can reduce compute demand.[161,162] Recent task-aware Transformer analysis for battery lifetime prediction further shows that hidden-dimension scaling and module-level pruning can reduce inference latency, FLOPs, and energy use while preserving, or in some cases improving, predictive accuracy.[163] Physics-guided inductive biases can also reduce the effective search space by embedding known constraints on charge conservation, transport, degradation, or thermal behavior.[149,164] More broadly, battery environmental impacts are strongly affected by capacity fade, usage conditions, manufacturing impacts, and electricity mix.[165] Computational sustainability could therefore be treated as a coupled technical and life-cycle metric.

Governance, accountability, and lifecycle responsibility.—Governance of BFMs may be best framed around life-cycle responsibility rather than generic AI oversight. BFMs may influence decisions in materials selection, manufacturing quality, BMS operation, second-life assessment, recycling, and safety management. Their outputs would therefore need to be supported by evidence on data provenance, validation conditions, uncertainty, and decision rationale. This aligns with emerging battery governance frameworks, including the EU Battery Regulation,[128] which emphasizes carbon footprint, recycled content, supply-chain due diligence, provenance, and life-cycle information.

Governance also determines who may participate in the development of BFMs and benefit from their outputs. High-value battery data are distributed across materials supply, manufacturing, operation, second-life use, and recycling, and selective data sharing may be needed to balance traceability, commercial confidentiality, accountability, and data governance requirements. If BFMs are developed mainly by data-rich organizations, they may reflect narrow chemistries, markets, or operating conditions. Practical governance may therefore combine openness with accountability through public benchmarks, standardized validation protocols, battery-passport-compatible metadata, uncertainty reporting, access control, and independent audits. Technical guidance for battery passport implementation further highlights the importance of interoperability, data quality, and system-level implementation requirements.[166] Therefore, the key challenge is to ensure that BFM outputs are safe and verifiable and that their benefits are broadly accessible across the battery life cycle, while avoiding the concentration of decision-making power in a few data-rich organizations.

Roadmap: Toward Lifecycle-Responsible BFMs

Translating BFMs from concept to deployment may require a staged pathway that links technical development with governance. The first step would be to establish curated, shareable battery datasets and standardized ontologies for interoperable data integration. On this basis, hybrid architectures that combine data-driven learning with electrochemical and thermo-mechanical constraints may improve reliability and transferability. Subsequent efforts could develop transferable and auditable prototypes validated under real-world operating conditions across multiple battery applications. Through this progression, BFMs may gradually support verifiable and traceable infrastructure for battery life-cycle analysis, spanning materials development, manufacturing, operation, and EOL management (Fig. 5).

Stage I: Curated Shared Datasets.

Curated, standardized, and shareable datasets may provide a foundational step for BFMs. Existing initiatives, including the Battery Data Genome,[23] automated battery-data platforms,[167] standardized data representations,[148] findable, accessible, interoperable, and reusable (FAIR)-oriented data-sharing ecosystems,[168] battery ontology frameworks,[169] and battery passports,[166] have already emphasized the need for interoperable battery data. However, broad adoption remains limited.

The main barriers are institutional as well as technical. Academic groups often receive limited credit for data curation, whereas companies may treat manufacturing parameters, quality control records, and field telemetry as sensitive assets. Data ownership, permitted use, liability, access control, and benefit sharing also remain insufficiently defined. Heterogeneous protocols, instruments, formats, sampling rates, and metadata conventions make even public datasets difficult to compare or reuse for pretraining. Therefore, for curated shared datasets to support BFMs in practice, future efforts may need to move beyond voluntary data release and establish mechanisms that make data sharing useful, safe, and creditable. Incentives could include data citation, dataset DOIs, benchmark recognition, FAIR data requirements, and contribution credit for high-quality data curation. Governance mechanisms could define tiered access, role-based permissions, data use agreements, anonymization procedures, intellectual property protection, and liability allocation. Technical standards could include metadata schemas, controlled vocabularies, battery ontologies, quality control protocols, and machine-readable formats. With these mechanisms, shared datasets are more likely

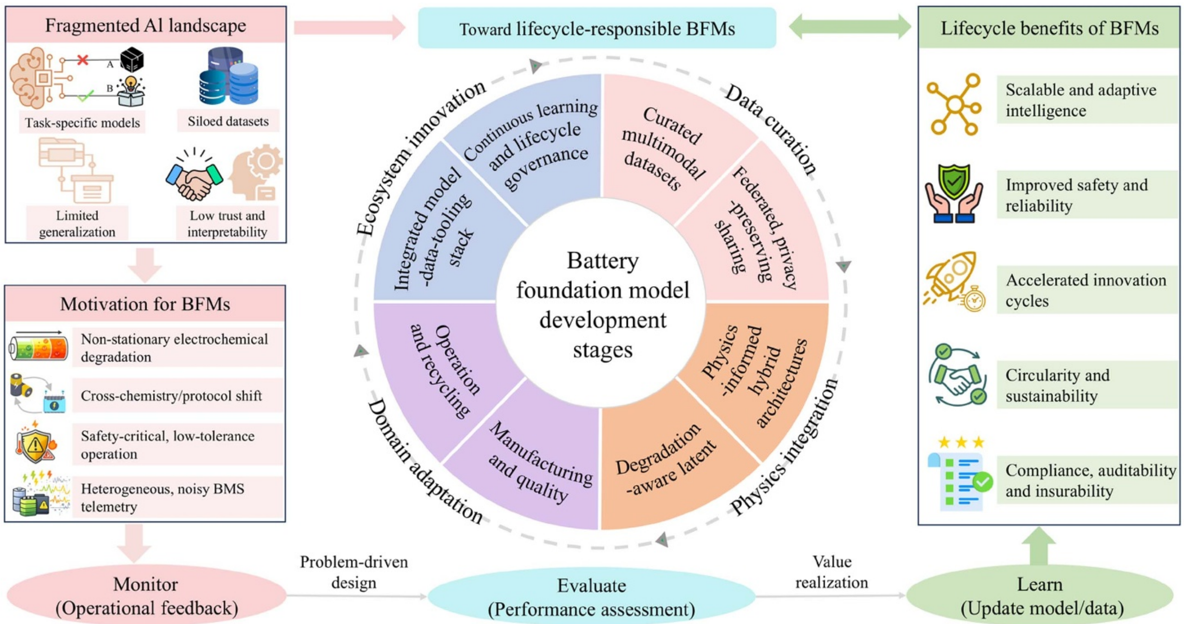

Figure 5. Roadmap toward lifecycle-responsible battery foundation models. By curating multimodal data and enabling privacy-preserving federated sharing, BFMs could combine physics-constrained hybrid architectures with degradation-aware representations to support domain adaptation across manufacturing, operation, diagnostics, and recycling. Supported by an integrated model–data–tooling stack and sustained lifecycle governance, a monitor–evaluate–learn feedback loop could support continual refinement and contribute to scalable intelligence, improved safety and reliability, accelerated innovation, enhanced circularity, and certification-oriented deployment with improved auditability and clearer insurability assessment.

to be adopted when data sharing is linked to compliance, certification, benchmarking, industrial value, and life-cycle traceability, rather than treated only as a voluntary scientific norm.

Stage II: Hybrid Models with Physics Priors.
As curated datasets become available, a next step would be to develop hybrid models that integrate data-driven learning with physical priors. Pretraining may improve representation learning under fragmented and privacy-constrained data, as suggested by self-supervised capacity estimation frameworks that reduce error under manufacturer-specific and aging-stage domain shifts. Reported gains include error reductions of up to 31.9% relative to the strongest baseline.[170] Pretraining alone, however, remains insufficient for safety-critical deployment. Coupling self-supervised learning (SSL) with mechanistic and physics-informed models, such as porous electrode theory,[171] xPatch,[172] and coupled thermo-electrochemical formulations,[173] could help anchor BFM behavior to core physical principles, including charge conservation, reaction kinetics, and heat dissipation. This integration may reduce the risk of physically implausible extrapolations.

Hybrid paradigms, such as physics-guided neural networks [174] and digital-twin frameworks,[175] provide practical and scalable pathways for this integration. By combining high-dimensional representation learning with governing equations, boundary conditions, and conservation constraints, these approaches may help restrict BFM predictions to physically admissible regimes. With embedded physical priors, BFMs could support more reliable interpolation within observed operating regimes and more principled extrapolation across chemistries, form factors, and usage patterns. Such hybridization could provide a bridge between large-scale statistical learning and physical science, supporting BFMs that are more adaptive and better suited to long-term deployment in safety-critical battery systems.

Stage III: Domain-Adapted BFMs.
The value of BFMs may depend less on ever-larger generic pretraining than on adaptation to concrete application contexts. Physical-prior-enriched pretraining could provide a reusable backbone, but deployable performance would require domain-specific refinement and validation. In manufacturing, industrial foundation-model systems such as MetaIndux illustrate how model training, domain adaptation, and application validation may support life-cycle-relevant battery applications.[176] During operation, industrial large models enhanced by domain knowledge could integrate online measurements with degradation mechanisms to support SOH estimation, anomaly detection, and cross-scenario generalization. At the EOL stage, life-cycle and cost comparisons of pyrometallurgical, hydrometallurgical, and direct cathode recycling show that direct cathode recycling may offer stronger environmental and economic benefits, providing quantitative support for policies that prioritize high-efficiency recycling pathways.[177]

Transfer learning, few-shot fine-tuning, and continual learning may support the application adaptation phase of BFMs. By combining general pre-trained knowledge with scenario-specific data, BFMs may improve task-specific performance while reducing the need for repeated training from scratch. Aligning adaptation workflows with regulatory and certification requirements may also improve the interpretability and auditability of model outputs. Through this process, BFMs may gradually move from general scientific models toward application-oriented systems embedded in manufacturing, vehicle operation, and recycling workflows. The adaptation layer may therefore help translate BFMs from academic prototypes into practical tools for battery life-cycle applications.

Stage IV: Integrated BFM Ecosystem.
A mature stage could be envisioned as an integrated ecosystem in which BFMs may serve as a shared intelligence layer connecting research, industrial practice, and governance. On the industrial side, existing cloud-based digital-twin BMSs have combined internet of things (IoT) data, ECMs, and SOC and SOH estimation algorithms to enable battery state diagnosis and aging monitoring, with system stability validated in both stationary and mobile applications.[178] On the academic front, the time-series foundation model Battery-Timer

leverages degradation-aware fine-tuning and knowledge distillation to achieve high-accuracy capacity prediction across scales and operating regimes, while substantially reducing deployment-time computational overhead.[179] From a governance perspective, when combined with unified data standards, enterprise data integration, and third-party audits, such platforms could enable continuously updated, verifiable, and auditable life-cycle-wide compliance oversight.[180,181] Realizing this ecosystem would require coordination across sectors and national boundaries. No single actor is likely to provide all necessary data, expertise, and computational resources. Public-private partnerships, international data alliances, and open governance models may therefore be needed to support trust, interoperability, and equitable access. The experience of climate science in building shared global infrastructures for data exchange suggests that the battery community could similarly establish mechanisms for collective stewardship and co-governance of BFMs. At its core, a unified BFM ecosystem could help reframe the role of AI in energy storage, shifting it from task-specific analytics toward a decision-support layer that connects multiple stages of the battery life cycle. By balancing innovation, regulatory compliance, and sustainability goals, such an ecosystem could help align technical advances with broader expectations for responsible and inclusive battery development.

Conclusions

AI in battery science appears to be moving from task-level perception and prediction toward reasoning-assisted decision support. Current applications, including cycle life prediction, fast-charging optimization, state estimation, anomaly detection, image-based characterization, and failure mode classification, can extract useful patterns from existing data, but remain limited in causal reasoning, cross-domain generalization, and autonomous decision-making. Whether AI can become decision-making infrastructure depends on the transition from fragmented, task-specific, and often single-modal datasets to standardized, multimodal, and life-cycle-wide data ecosystems. Future battery data ecosystems would need to integrate materials descriptors, cell designs, manufacturing parameters, electrochemical measurements, imaging and spectroscopy, field operation records, degradation histories, failure analysis, recycling information, and certification-relevant metadata. Data provenance, protocol standardization, uncertainty labels, and quality metrics are needed to distinguish electrochemical relationships from experimental artifacts or dataset-specific correlations. BFMs may therefore need to be developed together with these data ecosystems, relying not only on larger datasets but also on better structured, physically annotated, and continuously updated data streams. Realizing this vision requires collaboration among AI researchers, electrochemists, engineers, data scientists, manufacturers, and policymakers to establish shared data standards, hybrid modeling frameworks, governance mechanisms, and certification pathways. If successfully developed and rigorously validated, BFMs may help reshape how batteries are modeled, designed, manufactured, deployed, insured, certified, and recycled, supporting more integrated and life-cycle-responsible battery system management.

Acknowledgments

The authors acknowledge the support provided by the California Department of Transportation (Caltrans) and the U.S. Department of Transportation (USDOT) for the research project "Battery Technologies for Current and Future Electric Vehicles."

ORCID

Jingyuan Zhao https://orcid.org/0000-0002-6527-4294
Misheng Cai https://orcid.org/0009-0003-4212-8010
Yuqi Li https://orcid.org/0000-0003-1501-1549
Shengyu Tao https://orcid.org/0000-0001-5249-432X
Andrew F. Burke https://orcid.org/0000-0002-4245-0056
Stephen Harris https://orcid.org/0000-0002-5211-3934

References

1. W. Guo, S. B. Vilsen, Y. Li, A. Verma, D. I. Stroe and D. Brandell, "Uncovering the impact of battery design parameters on health and lifetime using short charging segments". *Energy Environ. Sci.*, **18**, 8462 (2025).
2. S. Zhu, B. Ramsundar, E. Annevelink, H. Lin, A. Dave, P. W. Guan and V. Viswanathan, "Differentiable modeling and optimization of non-aqueous Li-based battery electrolyte solutions using geometric deep learning". *Nat. Commun.*, **15**, 8649 (2024).
3. P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y. H. Liao and W. C. Chueh, "Closed-loop optimization of fast-charging protocols for batteries with machine learning". *Nature*, **578**, 397 (2020).
4. R. Cao, Z. Zhang, R. Shi, J. Lu, Y. Zheng, Y. Sun and S. Yang, "Model-constrained deep learning for online fault diagnosis in Li-ion batteries over stochastic conditions". *Nat. Commun.*, **16**, 1651 (2025).
5. R. K. Daniels, V. Kumar and A. Prabhakar, "A comparative study of data-driven thermal fault prediction using machine learning algorithms in air-cooled cylindrical Li-ion battery modules". *Renew. Sustain. Energy Rev.*, **207**, 114925 (2025).
6. T. Lombardo, M. Duquesnoy, H. El-bouysidy, F. Årén, A. Gallo-bueno, P. B. Jørgensen and A. A. Franco, "Artificial intelligence applied to battery research: hype or reality?". *Chem. Rev.*, **122**, 10899 (2021).
7. Q. Wang, B. Mao, S. I. Stoliarov and J. Sun, "A review of lithium-ion battery failure mechanisms and fire prevention strategies." *Prog. Energy Combust. Sci.*, **73**, 95 (2019).
8. J. Zhao and Z. Wang, "Specialized convolutional transformer networks for estimating battery health via transfer learning". *Energy Storage Mater.*, **71**, 103668 (2024).
9. Z. Lv and J. Zhao, "Resource-efficient artificial intelligence for battery capacity estimation using convolutional FlashAttention fusion networks". *eTransportation*, **23**, 100383 (2025).
10. R. Cao, Y. Li, J. Lu, F. Tian, H. Zhang, Y. Lu and S. Yang, "Multi-grouping-compatible frame-level unsupervised fault diagnosis and localization for electric vehicle battery packs in realistic conditions." *eTransportation*, **28**, 100579 (2026).
11. J. Zhao, X. Feng, M. K. Tran, M. Fowler, M. Ouyang and A. F. Burke, "Battery safety: fault diagnosis from laboratory to real world". *J. Power Sources*, **598**, 234111 (2024).
12. X. Feng, D. Ren, X. He and M. Ouyang, "Mitigating thermal runaway of lithium-ion batteries". *Joule*, **4**, 743 (2020).
13. M. Zhang, L. Liu, A. Stefanopoulou, J. Siegel, L. Lu, X. He and M. Ouyang, "Fusing phenomenon of lithium-ion battery internal short circuit". *J. Electrochem. Soc.*, **164**, A2738 (2017).
14. S. Zhou, R. Liu, Y. Wang, J. Lu and B. Jiang "BatteryBERT for realistic battery fault detection using point-masked signal modeling." In *2025 7th Int. Conf. on Industrial AI (IAI)*, p. 1 (2025). IEEE, Piscataway, NJ.
15. Y. Qin, S. Hu, Y. Lin, W. Chen, N. Ding, G. Cui and M. Sun, "Tool learning with foundation models". *ACM Comput. Surv.*, **57**, 1 (2024).
16. S. Li and H. Tang "Multimodal alignment and fusion: a survey." (2024) (arXiv:2411.17040).
17. T. Kojima, S. S. Gu, M. Reid, Y. Matsuo and Y. Iwasawa, "Large language models are zero-shot reasoners." *Advances in Neural Information Processing Systems*, vol 35, p. 22199 (2022).
18. T. Webb, K. J. Holyoak and H. Lu, "Emergent analogical reasoning in large language models". *Nat. Hum. Behav.*, **7**, 1526 (2023).
19. Q. Dong, L. Li, D. Dai, C. Zheng, J. Ma, R. Li and Z. Sui "A survey on in-context learning." In *Proc. 2024 Conf. on Empirical Methods in Natural Language Processing*, p. 1107 (2024).
20. A. Panagopoulou, L. Xue, N. Yu, J. Li, D. Li, S. Joty and J. C. Niebles "X–instructblip: a framework for aligning image, 3d, audio, video to llms and its emergent cross-modal reasoning." In *European Conf. on Computer Vision*, p. 177 (2024). *Cham*, Springer, Switzerland.
21. C. Lyu, M. Wu, L. Wang, X. Huang, B. Liu, Z. Du and Z. Tu "Macaw-LLM: multimodal language modeling with image, audio, video, and text integration" (arXiv:2306.09093).
22. J. Zhao, Y. Wang, M. Cai, L. Xie and H. Wang, "Large language models for battery prognostics". *J. Energy Chem.*, **115**, 159 (2026).
23. L. Ward, S. Babinec, E. J. Dufek, D. A. Howey, V. Viswanathan, M. Aykol and L. Teo, "Principles of the battery data genome". *Joule*, **6**, 2253 (2022).
24. K. Kim, K. Cho, K. Lee, J. Yoon and J. I. Choi, "Battery field data and why it matters: foundations for real-world electric vehicles." *eTransportation*, **26**, 100494 (2025).
25. Y. Liu, B. Hou, M. Ahmed, Z. Mao, J. Feng and Z. Chen, "A hybrid deep learning approach for remaining useful life prediction of lithium-ion batteries based on discharging fragments". *Appl. Energy*, **358**, 122555 (2024).
26. X. Liu, S. Tao, S. Fu, R. Ma, T. Cao, H. Fan and Y. Sun, "Binary multi-frequency signal for accurate and rapid electrochemical impedance spectroscopy acquisition in lithium-ion batteries". *Appl. Energy*, **364**, 123221 (2024).
27. F. Mansfeld, Z. Sun, C. H. Hsu and A. Nagiub, "Concerning trend removal in electrochemical noise measurements". *Corros. Sci.*, **43**, 341 (2001).
28. D. H. J. Baert and A. A. K. Vervaet, "Small bandwidth measurement of the noise voltage of batteries". *J. Power Sources*, **114**, 357 (2003).

29. W. Li, D. M. Lutz, L. Wang, K. J. Takeuchi, A. C. Marschilok and E. S. Takeuchi, "Peering into batteries: electrochemical insight through in situ and operando methods over multiple length scales". *Joule*, **5**, 77 (2021).
30. D. Liu, Z. Shadike, R. Lin, K. Qian, H. Li, K. Li and B. Li, "Review of recent development of in situ/operando characterization techniques for lithium battery research". *Adv. Mater.*, **31**, 1806620 (2019).
31. F. Lin, Y. Liu, X. Yu, L. Cheng, A. Singer, O. G. Shpyrko and M. M. Doeff, "Synchrotron X-ray analytical techniques for studying materials electrochemistry in rechargeable batteries". *Chem. Rev.*, **117**, 13123 (2017).
32. V. Wood, "X-ray tomography for battery research and development". *Nat. Rev. Mater.*, **3**, 293 (2018).
33. J. T. Lang, D. Kulkarni, C. W. Foster, Y. Huang, M. A. Sepe, S. Shimpalee and I. V. Zenyuk, "X-ray tomography applied to electrochemical devices and electrocatalysis". *Chem. Rev.*, **123**, 9880 (2023).
34. T. Sui, B. Song, J. Dluhos, L. Lu and A. M. Korsunsky, "Nanoscale chemical mapping of Li-ion battery cathode material by FIB-SEM and TOF-SIMS multi-modal microscopy". *Nano Energy*, **17**, 254 (2015).
35. D. J. Miller, C. Proff, J. G. Wen, D. P. Abraham and J. Bareño, "Observation of microstructural evolution in Li battery cathode oxide particles by in situ electron microscopy". *Adv. Energy Mater.*, **3**, 1098 (2013).
36. G. Feng, C. R. Rindt and S. G. Ritchie, "From diesel to electric: potential of drayage trucks transition in Southern California". *iscience*, **28**, 112629 (2025).
37. H. Cai, X. Jia, A. S. Chiu, X. Hu and M. Xu, "Siting public electric vehicle charging stations in Beijing using big-data informed travel patterns of the taxi fleet". *Transp. Res. Part D Transp. Environ.*, **33**, 39 (2014).
38. J. C. Kelly, J. S. MacDonald and G. A. Keoleian, "Time-dependent plug-in hybrid electric vehicle charging based on national driving patterns and demographics". *Appl. Energy*, **94**, 395 (2012).
39. A. Alonso-villar, B. Davíðsdóttir, H. Stefansson, E. I. Asgeirsson and R. Kristjansson, "Electrification potential for heavy-duty vehicles in harsh climate conditions: a case study based technical feasibility assessment". *J. Clean. Prod.*, **417**, 137997 (2023).
40. Y. Zhao, Z. Jiang, X. Chen, P. Liu, T. Peng and Z. Shu, "Toward environmental sustainability: data-driven analysis of energy use patterns and load profiles for urban electric vehicle fleets". *Energy*, **285**, 129465 (2023).
41. Y. Zhong, Y. Leng, Z. Gu, S. Guo, P. Li, S. Das, Y. Liu, J. Wan and Y. Liu, "Breaking interdisciplinary barriers in solid-state battery research: batteryAgent for multifaceted analysis". *J. Mater. Chem. A*, **13**, 37031 (2025).
42. Z. Wu, L. Qiu, A. Ross, E. Akyürek, B. Chen, B. Wang and Y. Kim "Reasoning or reciting? exploring the capabilities and limitations of language models through counterfactual tasks." In *Proc. 2024 Conf. North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)* p. 1819 (2024).
43. K. Mahowald, A. A. Ivanova, I. A. Blank, N. Kanwisher, J. B. Tenenbaum and E. Fedorenko, "Dissociating language and thought in large language models". *Trends Cogn. Sci.*, **28**, 517 (2024).
44. E. Yee, A. Li, C. Tang, Y. H. Jung, R. Paturi and L. Bergen "Dissociation of faithful and unfaithful reasoning in LLMs." (2024) (arXiv:2405.15092).
45. Z. Bai, P. Wang, T. Xiao, T. He, Z. Han, Z. Zhang and M. Z. Shou "Hallucination of multimodal large language models: a survey." (2024). (arXiv:2404.18930).
46. X. Feng, M. Ouyang, X. Liu, L. Lu, Y. Xia and X. He, "Thermal runaway mechanism of lithium-ion battery for electric vehicles: a review". *Energy Storage Mater.*, **10**, 246 (2018).
47. J. Zhao, H. Ling, J. Liu, J. Wang, A. F. Burke and Y. Lian, "Machine learning for predicting battery capacity for electric vehicles". *eTransportation*, **15**, 100214 (2023).
48. H. Peng, C. Liu and H. Li, "Large-language-model-enabled health management for internet of batteries in electric vehicles". *IEEE Internet Things J.*, **12**, 6082 (2024).
49. D. A. Boiko, R. MacKnight, B. Kline and G. Gomes, "Autonomous chemical research with large language models". *Nature*, **624**, 570 (2023).
50. C. Bian, Z. Duan, Y. Hao, S. Yang and J. Feng, "Exploring large language model for generic and robust state-of-charge estimation of Li-ion batteries: a mixed prompt learning method". *Energy*, **302**, 131856 (2024).
51. L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin and R. Lowe, "Training language models to follow instructions with human feedback." *Adv. Neural Inf. Process. Syst.*, **35**, 27730 (2022).
52. J. Wei, M. Bosma, V. Y. Zhao, K. Guu, A. W. Yu, B. Lester and Q. V. Le "Finetuned language models are zero-shot learners." (2021). (arXiv:2109.01652).
53. P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal and D. Kiela, "Retrieval-augmented generation for knowledge-intensive nlp tasks." *Adv. Neural Inf. Process. Syst.*, **33**, 9459 (2020).
54. E. J. Hu, Y. Shen, P. Wallis, Z. Allen-zhu, Y. Li, S. Wang and W. Chen, "Lora: low-rank adaptation of large language models." *Iclr*, **1**, 3 (2022).
55. B. Lyu, Y. Cao, D. Watson-parris, L. Bergen, T. Berg-kirkpatrick and R. Yu "Adapting while learning: grounding llms for scientific problems with intelligent tool usage adaptation." (2024). (arXiv:2411.00412).
56. Y. Zong, O. Mac Aodha and T. M. Hospedales, "Self-supervised multimodal learning: a survey". *IEEE Trans. Pattern Anal. Mach. Intell.*, **47**, 5299 (2024).
57. X. Yang and A. L. Rogach, "Electrochemical techniques in battery research: a tutorial for nonelectrochemists". *Adv. Energy Mater.*, **9**, 1900747 (2019).
58. F. Tao, B. Xiao, Q. Qi, J. Cheng and P. Ji, "Digital twin modeling". *J. Manuf. Syst.*, **64**, 372 (2022).
59. Y. Che, J. Guo, Y. Zheng, D. I. Stroe, W. Liu, X. Hu and R. Teodorescu, "Unlocking interpretable prediction of battery random discharge capacity with domain adaptative physics constraint". *Adv. Energy Mater.*, **15**, 2405506 (2025).
60. H. Liu, C. Li, X. Hu, J. Li, K. Zhang, Y. Xie, R. Wu and Z. Song, "Multimodal framework for battery state of health evaluation using open-source electric vehicle data". *Nat. Commun.*, **16**, 1137 (2025).
61. G. Liang, C. Zhang, L. Yang, Y. Liu, M. Liu, X. Xiong and R. Che, "Probing interfacial nanostructures of electrochemical energy storage systems by in-situ transmission electron microscopy". *Nano-Micro Lett.*, **17**, 245 (2025).
62. B. K. Cho, S. Y. Jung, S. J. Park, J. H. Hyun and S. H. Yu, "In situ/operando imaging techniques for next-generation battery analysis". *ACS Energy Lett.*, **9**, 4068 (2024).
63. M. B. Dixit, J. S. Park, P. Kenesei, J. Almer and K. B. Hatzell, "Status and prospect of in situ and operando characterization of solid-state batteries". *Energy Environ. Sci.*, **14**, 4672 (2021).
64. M. G. Boebinger, J. A. Lewis, S. E. Sandoval and M. T. McDowell, "Understanding transformations in battery materials using in situ and operando experiments: progress and outlook". *ACS Energy Lett.*, **5**, 335 (2020).
65. C. Zhang, Z. Yang, X. He and L. Deng, "Multimodal intelligence: representation learning, information fusion, and applications". *IEEE J. Sel. Top. Signal Process.*, **14**, 478 (2020).
66. T. Baltrušaitis, C. Ahuja and L. P. Morency, "Multimodal machine learning: a survey and taxonomy". *IEEE Trans. Pattern Anal. Mach. Intell.*, **41**, 423 (2018).
67. R. Wu, H. Wang, H. T. Chen and G. Carneiro "Deep multimodal learning with missing modality: a survey." (2024). (arXiv:2409.07825).
68. Q. Wang, M. Ye, X. Cai, D. U. Sauer and W. Li, "Transferable data-driven capacity estimation for lithium-ion batteries with deep learning: a case study from laboratory to field applications". *Appl. Energy*, **350**, 121747 (2023).
69. F. Wang, Z. Zhai, Z. Zhao, Y. Di and X. Chen, "Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis". *Nat. Commun.*, **15**, 4332 (2024).
70. T. Hofmann, J. Hamar, M. Rogge, C. Zoerr, S. Erhard and J. Philipp Schmidt, "Physics-informed neural networks for state of health estimation in lithium-ion batteries". *J. Electrochem. Soc.*, **170**, 090524 (2023).
71. S. Cai, Z. Wang, S. Wang, P. Perdikaris and G. E. Karniadakis, "Physics-informed neural networks for heat transfer problems". *J. Heat Transfer*, **143**, 060801 (2021).
72. Y. Ji and D. T. Schwartz, "Second-harmonic nonlinear electrochemical impedance spectroscopy: part I. Analytical theory and equivalent circuit representations for planar and porous electrodes". *J. Electrochem. Soc.*, **170**, 123511 (2023).
73. H. Pang, X. Yan, N. Jiang, X. Qu, B. Jing, A. F. Burke and J. Zhao, "Electro-thermal-aging modeling of Li-ion batteries with active-material loss." *Appl. Energy*, **410**, 127539 (2026).
74. K. Xu and C. Wang, "Batteries: widening voltage windows". *Nat. Energy*, **1**, 1 (2016).
75. Q. Li, H. Liu, Y. Ye, K. J. Li, F. Wu, L. Li and R. Chen, "The critical importance of stack pressure in batteries". *Nat. Energy*, **10**, 1064 (2025).
76. Z. Rao, P. Lyu, M. Li, X. Liu and X. Feng, "A thermal perspective on battery safety". *Nat. Rev. Clean Technol.*, **1**, 511 (2025).
77. M. A. Nabian and H. Meidani, "Physics-driven regularization of deep neural networks for enhanced engineering design and analysis". *J. Comput. Inf. Sci. Eng.*, **20**, 011006 (2020).
78. S. Park, H. Lee, S. Choi, J. Lim, S. Kim, J. Song and Y. M. Lee, "Microstructure-based digital twin thermo-electrochemical modeling of LIBs at the cell-to-module scale". *eTransportation*, **22**, 100370 (2024).
79. J. Zhao, Y. Hong, Y. Zheng, Y. Li, X. Feng, M. Fowler, M. Ouyang, A. F. Burke and D. Sperling, "Battery system integration for electric mobility". *Joule*, **10**, 102445 (2026).
80. V. Sulzer, P. Mohtat, A. Aitio, S. Lee, Y. T. Yeh, F. Steinbacher and D. A. Howey, "The challenge and opportunity of battery lifetime prediction from field data". *Joule*, **5**, 1934 (2021).
81. H. Pang, X. Qu, K. Chen, X. Yan, A. F. Burke and J. Zhao, "A transferable deep learning framework for cross-chemistry and cross-scale state of health estimation in lithium-ion batteries." *Energy*, **354**, 141079 (2026).
82. G. Dos Reis, C. Strange, M. Yadav and S. Li, "Lithium-ion battery data and where to find it". *Energy AI*, **5**, 100081 (2021).
83. J. Stempel "BASF sues Berkshire's Duracell for allegedly stealing battery secrets." *Reuters* (2025). (available at: www.reuters.com) (Accessed June 5, 2026).
84. L. Cheng, R. S. Assary, X. Qu, A. Jain, S. P. Ong, N. N. Rajput and L. A. Curtiss, "Accelerating electrolyte discovery for energy storage with high-throughput screening". *J. Phys. Chem. Lett.*, **6**, 283 (2015).
85. Z. Han, S. Tao, Y. Jia, M. Zhang, R. Ma, X. Xiao and G. Zhou, "Data-driven insight into the universal structure–property relationship of catalysts in lithium–sulfur batteries". *J. Am. Chem. Soc.*, **147**, 22851 (2025).
86. Z. Han, R. Gao, T. Wang, S. Tao, Y. Jia, Z. Lao and G. Zhou, "Machine-learning-assisted design of a binary descriptor to decipher electronic and structural effects on sulfur reduction kinetics". *Nat. Catal.*, **6**, 1073 (2023).
87. N. He, Q. Wang, Z. Lu, Y. Chai and F. Yang, "Early prediction of battery lifetime based on graphical features and convolutional neural networks". *Appl. Energy*, **353**, 122048 (2024).
88. G. Lai, J. Lin, W. Mo, X. Li, X. Jin, M. Wu and F. Liu, "Expanding the electrochemical stability window: unraveling the role of solvent polarity and a WiSE-compatible strategy". *J. Am. Chem. Soc.*, **147**, 13071 (2025).

89. E. C. Y. Yuan, Y. Liu, J. Chen, P. Zhong, S. Raja, T. Kreiman and T. Head-gordon, "Foundation models for atomistic simulation of chemistry and materials". *Nat. Rev. Chem.*, **10**, 212 (2026).
90. X. Du, J. Meng, Z. Xue, Y. Amirat, F. Gao and M. Benbouzid, "Revolutionizing battery safety: real-time insights with dynamic electrochemical impedance spectroscopy". *ACS Energy Lett.*, **10**, 2292 (2025).
91. X. Wang, Y. Li and Y. S. Meng, "Cryogenic electron microscopy for characterizing and diagnosing batteries". *Joule*, **2**, 2225 (2018).
92. B. Sanchez-lengeling and A. Aspuru-guzik, "Inverse molecular design using machine learning: generative models for matter engineering". *Science*, **361**, 360 (2018).
93. D. M. Anstine and O. Isayev, "Generative models as an emerging paradigm in the chemical sciences". *J. Am. Chem. Soc.*, **145**, 8736 (2023).
94. V. W. Hu and D. T. Schwartz, "Low error estimation of half-cell thermodynamic parameters from whole-cell Li-ion battery experiments: physics-based model formulation, experimental demonstration, and an open software tool". *J. Electrochem. Soc.*, **169**, 030539 (2022).
95. J. M. Reniers, G. Mulder and D. A. Howey, "Review and performance comparison of mechanical-chemical degradation models for lithium-ion batteries". *J. Electrochem. Soc.*, **166**, A3189 (2019).
96. A. Nekahi, E. Feyzi, M. Srivastava, F. Yeganehdoust, A. K. M. R. Reddy and K. Zaghib, "Advanced lithium-ion battery process manufacturing equipment for gigafactories: past, present, and future perspectives". *iscience*, **28**, 112691 (2025).
97. R. Zhan, D. Ren, S. Liu, Z. Chen, X. Liu, W. Wang and Y. Sun, "A paradigm of calendaring-driven electrode microstructure for balanced battery energy density and power density". *Adv. Energy Mater.*, **13**, 2202544 (2023).
98. S. Liu, Z. Zhou, J. Jiang, S. Chen, X. Li and X. G. Yang, "Ultra-rapid electrolyte leakage diagnosis for lithium-ion batteries based on ultrasonic guided waves". *J. Energy Storage*, **123**, 116751 (2025).
99. M. D. Murbach and D. T. Schwartz, "Analysis of li-ion battery electrochemical impedance spectroscopy data: an easy-to-implement approach for physics-based parameter estimation using an open-source tool". *J. Electrochem. Soc.*, **165**, A297 (2018).
100. J. Li, E. Hu, W. Wei and F. Shi, "From wafers to electrodes: transferring automatic optical inspection (AOI) for multiscale characterization of smart battery manufacturing." *Adv. Funct. Mater.*, **36**, e28142 (2026).
101. O. Özgül and A. Sammler Optimizing electrode production with real-time viscosity measurement." (2025), (available at: https://analyzing-testing.netzsch.com). (Accessed December 20, 2025).
102. Capgemini. "How to accelerate EV battery manufacturing in gigafactories." (2024), (available at:www.capgemini.com). (Accessed December 20, 2025).
103. H. Jing, S. Ou, Z. Lv, H. Guo, A. F. Burke and J. Zhao, "Battery safety: mechanisms, monitoring, and machine intelligence." *Adv. Energy Mater.*, **16**, e71093 (2026).
104. J. Zhao, X. Feng, Q. Pang, M. Fowler, Y. Lian, M. Ouyang and A. F. Burke, "Battery safety: machine learning-based prognostics." *Prog. Energy Combust. Sci.*, **102**, 101142 (2024).
105. D. Fioriti, C. Scarpelli, L. Pellegrino, G. Lutzemberger, E. Micolano and S. Salamone, "Battery lifetime of electric vehicles by novel rainflow-counting algorithm with temperature and C-rate dynamics: effects of fast charging, user habits, vehicle-to-grid and climate zones". *J. Energy Storage*, **59**, 106458 (2023).
106. A. Maheshwari, N. G. Paterakis, M. Santarelli and M. Gibescu, "Optimizing the operation of energy storage using a non-linear lithium-ion battery degradation model". *Appl. Energy*, **261**, 114360 (2020).
107. M. T. Von Srbik, M. Marinescu, R. F. Martinez-botas and G. J. Offer, "A physically meaningful equivalent circuit network model of a lithium-ion battery accounting for local electrochemical and thermal behavior, variable double layer capacitance and degradation". *J. Power Sources*, **325**, 171 (2016).
108. C. Xu, E. Zhang, K. Jiang and K. Wang, "Dual fuzzy-based adaptive extended Kalman filter for state of charge estimation of liquid metal battery". *Appl. Energy*, **327**, 120091 (2022).
109. F. Li, Y. Min, Y. Zhang, Y. Zhang, H. Zuo and F. Bai, "State-of-health estimation method for fast-charging lithium-ion batteries based on stacking ensemble sparse Gaussian process regression". *Reliab. Eng. Syst. Saf.*, **242**, 109787 (2024).
110. J. Zhao, Z. Lv, Y. Che, Y. Li and A. F. Burke, "End-to-end millisecond-level battery aging diagnostics for electric vehicles". *J. Energy Storage*, **143**, 119638 (2026).
111. Z. Lv, S. S. Ou, H. Jing, H. Dong, D. Shi and J. Zhao, "Battery state-of-health estimation via thermoelectric graph learning and interpretable temporal decomposition for real-world electric vehicles". *J. Energy Storage*, **153**, 120956 (2026).
112. T. Bond, R. Gauthier, S. Gasilov and J. R. Dahn, "In-situ computed tomography of particle microcracking and electrode damage in cycled NMC622/graphite pouch cell batteries". *J. Electrochem. Soc.*, **169**, 080531 (2022).
113. Y. Chen, Y. Yao, Z. Yao, W. Li, X. Shen, J. Song and K. Wu, "Fracture behavior of NCM polycrystalline particles in lithium-ion batteries under extreme conditions". *Nano Energy*, **141**, 111104 (2025).
114. P. P. Paul, V. Thampy, C. Cao, H. G. Steinrück, T. R. Tanim, A. R. Dunlop and J. N. Weker, "Quantification of heterogeneous, irreversible lithium plating in extreme fast charging of lithium-ion batteries". *Energy Environ. Sci.*, **14**, 4979 (2021).
115. L. Hartmann, L. Reuter, L. Wallisch, A. Beiersdorfer, A. Adam, D. Goldbach and J. Wandt, "Depletion of electrolyte salt upon calendaric aging of lithium-ion batteries and its effect on cell performance". *J. Electrochem. Soc.*, **171**, 060506 (2024).
116. X. Yu, R. Chen, L. Gan, H. Li and L. Chen, "Battery safety: from lithium-ion to solid-state batteries". *Engineering*, **21**, 9 (2023).
117. L. Liu, X. Feng, C. Rahe, W. Li, L. Lu, X. He and M. Ouyang, "Internal short circuit evaluation and corresponding failure mode analysis for lithium-ion batteries". *J. Energy Chem.*, **61**, 269 (2021).
118. N. Fallah and C. Fitzpatrick, "How will retired electric vehicle batteries perform in grid-based second-life applications? A comparative techno-economic evaluation of used batteries in different scenarios". *J. Clean. Prod.*, **361**, 132281 (2022).
119. R. Ma, S. Tao, X. Sun, Y. Ren, C. Sun, G. Ji and G. Zhou, "Pathway decisions for reuse and recycling of retired lithium-ion batteries considering economic and environmental functions". *Nat. Commun.*, **15**, 7641 (2024).
120. S. Tao, R. Ma, Z. Zhao, G. Ma, L. Su, H. Chang and G. Zhou, "Generative learning assisted state-of-health estimation for sustainable battery recycling with random retirement conditions". *Nat. Commun.*, **15**, 10154 (2024).
121. S. Tao, H. Liu, C. Sun, H. Ji, G. Ji, Z. Han and H. Sun, "Collaborative and privacy-preserving retired battery sorting for profitable direct recycling via federated machine learning". *Nat. Commun.*, **14**, 8032 (2023).
122. S. Tao, R. Ma, Y. Chen, Z. Liang, H. Ji, Z. Han and G. Zhou, "Rapid and sustainable battery health diagnosis for recycling pretreatment using fast pulse test and random forest machine learning". *J. Power Sources*, **597**, 234156 (2024).
123. Y. Zang and Y. Wang "Robotic disassembly of electric vehicle batteries: an overview." In *2022 27th Int. Conf. Automation and Computing (ICAC)*, p. 1 (2022), IEEE, Piscataway, NJ.
124. C. Erdogan, C. A. Contreras, A. Rastegarpanah, M. Chiou and R. Stolkin "Intent-driven LLM ensemble planning for flexible multi-robot disassembly: demonstration on EV batteries." (2025). (arXiv:2510.17576).
125. Y. Jin, T. Zhang and M. Zhang, "Advances in intelligent regeneration of cathode materials for sustainable lithium-ion batteries". *Adv. Energy Mater.*, **12**, 2201526 (2022).
126. Y. Yao, M. Zhu, Z. Zhao, B. Tong, Y. Fan and Z. Hua, "Hydrometallurgical processes for recycling spent lithium-ion batteries: a critical review". *ACS Sustain. Chem. Eng.*, **6**, 13611 (2018).
127. O. S. Kwon and I. L. Sohn, "Fundamental thermokinetic study of a sustainable lithium-ion battery pyrometallurgical recycling process". *Resources, Conservation and Recycling*, **158**, 104809 (2020).
128. European Parliament and Council of the European Union. "Regulation (EU) 2023/1542 concerning batteries and waste batteries." (2023), (available at: https://eur-lex.europa.eu/eli/reg/2023/1542/oj). (Accessed June 6, 2026).
129. Y. Jiang and X. Zhang, "Product life cycle environmental regulation and corporate carbon emissions: quasi-natural experiment from extended producer responsibility system." *Energy Econ.*, **153**, 109066 (2025).
130. L. Adam, R. Pellerin and B. Agard, "Analysis of rule-based models for predicting energy consumption in electric bus fleets: a case study with telemetry data." *Autom. Robot. Commun. Ind.,*, **4**, 68 (2025).
131. S. Wang, F. Zhou, J. Zhao, H. Wang, M. Cai, H. Wang and H. Yin, "Mitigating overestimation in lithium-ion battery recycling LCA: the critical role of ex-post data and operational parameters." *Chem. Eng. J.*, **521**, 166261 (2025).
132. D. P. Finegan, M. Scheel, J. B. Robinson, B. Tjaden, I. Hunt, T. J. Mason and P. R. Shearing, "In-operando high-speed tomography of lithium-ion batteries during thermal runaway". *Nat. Commun.*, **6**, 6924 (2015).
133. W. B. Hawley and J. Li, "Electrode manufacturing for lithium-ion batteries—Analysis of current and next generation processing". *J. Energy Storage*, **25**, 100862 (2019).
134. A. Tabassum, S. Allu, R. Kannan and N. Muralidhar "Counter data paucity through adversarial invariance encoding: a case study on modeling battery thermal runaway." In *2024 IEEE Int. Conf. Big Data (BigData)* p. 2224 (2024). IEEE, Piscataway, NJ.
135. L. Li, Y. Fan, M. Tse and K. Y. Lin, "A review of applications in federated learning". *Comput. Ind. Eng.*, **149**, 106854 (2020).
136. H. Zhu, J. Xu, S. Liu and Y. Jin, "Federated learning on non-IID data: a survey". *Neurocomputing*, **465**, 371 (2021).
137. B. McMahan, E. Moore, D. Ramage, S. Hampson and B. A. Arcas, "Communication-efficient learning of deep networks from decentralized data." *AI and Statistics* p. 1273 PMLR, (2017).
138. L. Zhu, Z. Liu and S. Han, "Deep leakage from gradients." *Advances in Neural Information Processing Systems*, 32 (2019).
139. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang and L. Yang, "Physics-informed machine learning". *Nat. Rev. Phys.*, **3**, 422 (2021).
140. R. Tan, W. Hong, J. Tang, X. Lu, R. Ma, X. Zheng and T. Y. Zhang "BatteryLife: a comprehensive dataset and benchmark for battery life prediction." In *Proc.31st ACM SIGKDD Conf. Knowledge Discovery and Data Mining*, vol 2 p. 5789 (2025).
141. A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon and E. D. Cubuk, "Scaling deep learning for materials discovery". *Nature*, **624**, 80 (2023).
142. C. Chen, D. T. Nguyen, S. J. Lee, N. A. Baker, A. S. Karakoti, L. Lauw and D. Thien "Accelerating computational materials discovery with artificial intelligence and cloud high-performance computing: from large-scale screening to experimental validation." (arXiv:2401.04070).
143. Q. Du, M. M. Sullivan, J. E. Saal and F. Huber "AI-guided design and optimization of graphite-based anodes via iterative experimental feedback." (2026). (arXiv:2606.00187).
144. A. Yao, S. M. Benson and W. C. Chueh, "Critically assessing sodium-ion technology roadmaps and scenarios for techno-economic competitiveness against lithium-ion batteries". *Nat. Energy*, **10**, 404 (2025).

145. J. Zhao, Y. Che, Y. Li, H. Jing, S. Ou and D. Sperling, "Trade-offs in battery integration for transportation." *Nat. Rev. Clean Technol.* **2**, 605–17 (2026).
146. X. G. Yang, T. Liu, S. Ge, E. Rountree and C. Y. Wang, "Challenges and key requirements of batteries for electric vertical takeoff and landing aircraft". *Joule*, **5**, 1644 (2021).
147. Y. Wang, J. Bai, M. S. Eshaghi, C. Anitescu, X. Zhuang, T. Rabczuk and Y. Liu, "Transfer learning in physics-informed neurals networks: full fine-tuning, lightweight fine-tuning, and low-rank adaptation". *Int. J. Mech. Syst. Dyn.*, **5**, 212 (2025).
148. C. N. Lininger, T. Thai, T. R. Juran, E. S. Leland and T. Z. Sholklapper, "Voltaiq data format—A standard data format for collection of battery data to enable big data comparisons and analyses across the battery lifecycle". *Front. Energy Res.*, **10**, 1059154 (2022).
149. S. Navidi, A. Thelen, T. Li and C. Hu, "Physics-informed machine learning for battery degradation diagnostics: a comparison of state-of-the-art methods". *Energy Storage Mater.*, **68**, 103343 (2024).
150. S. Sripad, A. Bills and V. Viswanathan, "A review of safety considerations for batteries in aircraft with electric propulsion". *MRS Bull.*, **46**, 435 (2021).
151. R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti and D. Pedreschi, "A survey of methods for explaining black box models". *ACM Comput. Surv.*, **51**, 1 (2018).
152. J. Zhao, Y. Che, Y. Li and S. Harris, "From hype to impact: a roadmap for trustworthy battery AI." *eTransportation*, **28**, 100546 (2026).
153. RTCA. "DO-311A: minimum operational performance standards for rechargeable lithium batteries and battery systems. radio technical commission for aeronautics." (2017). (available at: www.rtca.org/). (Accessed June 6, 2026).
154. ISO. "ISO 26262: road vehicles—functional safety." International Organization for Standardization, (2018), (available at: www.iso.org/home.html). (Accessed June 6, 2026).
155. A. Zewe "Explained: generative AI's environmental impact." MIT News, (2025), (available at: https://news.mit.edu) (Accessed June 6, 2026).
156. P. Li, Y. Liu, J. Yang and S. Ren "Toward socially and environmentally responsible AI." (2024). (arXiv:2407.05176).
157. D. Patterson, J. Gonzalez, Q. Le, C. Liang, L. M. Munguia, D. Rothchild and J. Dean "Carbon emissions and large neural network training." (2021). (arXiv:2104.10350).
158. A. S. Luccioni, S. Viguier and A. L. Ligozat, "Estimating the carbon footprint of bloom, a 176b parameter language model." *J. Mach. Learn. Res.*, **24**, 1 (2023).
159. S. Luccioni, Y. Jernite and E. Strubell "Power hungry processing: watts driving the cost of AI deployment?." In *Proc. 2024 ACM Conf. Fairness, Accountability, and Transparency*, p. 85 (2024).
160. J. K. Barillas, J. Li, C. Günther and M. A. Danzer, "A comparative study and validation of state estimation algorithms for Li-ion batteries in battery management systems". *Appl. Energy*, **155**, 455 (2015).
161. Y. Sy, C. Cerisara and I. Illina "Lillama: large language models compression via low-rank feature distillation." (2024). (arXiv:2412.16719).
162. G. I. Kim, S. Hwang and B. Jang, "Efficient compressing and tuning methods for large language models: a systematic literature review." *ACM Comput. Surv.*, **57**, 1 (2025).
163. J. Zhao, M. Cai, Z. Wang, Y. Wang, Y. Li, B. Dong, H. Wang, A. F. Burke, S. Onori and S. Harris, "Performance–complexity trade-offs in battery lifetime prediction with task-aware transformers." *Adv. Sci.*, e24179 (2026).
164. Z. Yang, J. Nan, Z. Ge, M. Cai and J. Zhao, "Physics-informed dual-stream networks for battery health estimation." *Results Eng.*, **29**, 109973 (2026).
165. P. Marques, R. Garcia, L. Kulay and F. Freire, "Comparative life cycle assessment of lithium-ion batteries for electric vehicles addressing capacity fade". *J. Clean. Prod.*, **229**, 787 (2019).
166. Battery Passport Technical Guidance (2024), (available at: https://thebatterypass.eu) (Accessed June 6, 2026).
167. Voltaiq, "What is Voltaiq?." (2025), (available at: www.voltaiq.com/platform) (Accessed December 20, 2025).
168. I. E. Castelli, D. J. Arismendi-arrieta, A. Bhowmik, I. Cekic-laskovic, S. Clark, R. Dominko and T. Vegge, "Data management plans: the importance of data management in the BIG-MAP project". *Batter. Supercaps*, **4**, 1803 (2021).
169. BIG-MAP. BattINFO. Battery interface genome—materials acceleration platform (2025), (available at:www.big-map.eu/dissemination/battinfo?) (Accessed: 20 December, 2025).
170. N. Tvlg, "From sparse privacy-compliant data to accurate assessment: a self-supervised paradigm for EV battery capacity estimation." (available at: www.authorea.com/doi/full/10.22541/au.176038142.20913359/v1) (Accessed June 6, 2026).
171. M. D. Berliner, H. Zhao, S. Das, M. Forsuelo, B. Jiang, W. H. Chueh and R. D. Braatz, "Nonlinear identifiability analysis of the porous electrode theory model of lithium-ion batteries". *J. Electrochem. Soc.*, **168**, 090546 (2021).
172. Z. Yuan, Z. Deng, Y. He, Z. Ning and J. Liu, "Multi-step prediction of battery state of health based on self-supervised pre-training and transfer learning using the xPatch model." *Energy*, **341**, 139410 (2025).
173. W. B. Gu and C. Y. Wang, "Thermal-electrochemical modeling of battery systems". *J. Electrochem. Soc.*, **147**, 2910 (2000).
174. X. Dai, R. Zhou, J. Zhang, K. He, F. Lin and H. Ma, "SocNet: a physics-guided neural network for battery state-of-charge estimation robust to temperature variations and sensor noises". *IEEE Trans. Transp. Electr.*, **11**, 11165 (2025).
175. Y. Qin, A. Arunan and C. Yuen, "Digital twin for real-time Li-ion battery state of health estimation with partially discharged cycling data". *IEEE Trans. Ind. Inform.*, **19**, 7247 (2023).
176. L. Ren, H. Wang, J. Dong, Z. Jia, S. Li, Y. Wang and B. Li, "Industrial foundation model". *IEEE Trans. Cybern.*, **55**, 2286 (2025).
177. O. Velázquez-martínez, J. Valio, A. Santasalo-aarnio, M. Reuter and R. Serna-guerrero, "A critical review of lithium-ion battery recycling processes from a circular economy perspective". *Batteries*, **5**, 68 (2019).
178. W. Li, M. Rentemeister, J. Badeda, D. Jöst, D. Schulte and D. U. Sauer, "Digital twin for battery systems: cloud battery management system with online state-of-charge and state-of-health estimation". *J. Energy Storage*, **30**, 101557 (2020).
179. J. Chan, Z. Chen and E. Pan, "Foundation models knowledge distillation for battery capacity degradation forecast." *Reliab. Eng. Syst. Saf.*, **270**, 112139 (2025).
180. A. Weng, E. Dufek and A. Stefanopoulou, "Battery passports for promoting electric vehicle resale and repurposing". *Joule*, **7**, 837 (2023).
181. K. Berger, J. P. Schöggl and R. J. Baumgartner, "Digital battery passports to enable circular and sustainable value chains: conceptualization and use cases". *J. Clean. Prod.*, **353**, 131492 (2022).