STM JOURNALS

Journal of
Polymer & Composites

ISSN: 2321-2810 (Online)
ISSN: 2321-8525 (Print)
Volume 14, Special Issue 3, 2026
DOI (Journal): 10.37591/JoPC

https://journals.stmjournals.com/jopc

Review JoPC

AI-Based Discovery of High-Performance Energy Storage Polymer Composites: A Comprehensive Review

Manas Kumar Yogi[1]*, Darapu Uma[2], Yamuna Mundru[3], Manam Siva Sai[4]

Abstract

The accelerating global demand for high-performance energy storage systems has stimulated significant research into advanced polymer composites as next-generation electrolytes, electrode binders, and functional membranes for batteries, supercapacitors, and photovoltaic devices. However, the vast compositional and structural design space of polymer materials presents formidable challenges for conventional trial-and-error discovery strategies, which remain slow, costly, and biased by prior expert knowledge. Machine learning (ML) and artificial intelligence (AI) have emerged as transformative tools for navigating this complexity, enabling rapid prediction of electrochemical properties, de novo design of polymer electrolytes, and precise optimization of nanostructures for supercapacitors and batteries. This review systematically examines the application of ML techniques, including graph neural networks, Bayesian optimization, variational autoencoders, and transformer-based language models, for the discovery of energy storage polymer composites. The discussion critically evaluates ML-driven advancements across lithium-ion batteries, flexible energy storage devices, and solar energy materials, drawing on quantitative performance benchmarks reported in the primary literature. Emerging strategies such as active learning, multi-fidelity data fusion, physics-informed neural networks, and polymer-specific foundation models are discussed alongside persistent challenges related to data scarcity, model interpretability, and the translation gap between computational prediction and experimental synthesis. The review further addresses the landscape of open polymer property databases, the role of autonomous closed-loop experimentation in accelerating materials discovery, and the importance of reproducible, well-documented machine learning pipelines for the field to mature beyond proof-of-concept demonstrations. By consolidating evidence from verified primary sources and presenting original comparative analyses across methods and application domains, this review provides researchers, materials scientists, and computational chemists with an actionable, evidence-based perspective on the current state and future trajectory of AI-accelerated, sustainable energy storage polymer composite discovery.

***Author for Correspondence**
Manas Kumar Yogi

[1]Assistant Professor, Department of Computer Science and Engineering, Pragati Engineering College (Autonomous), Surampalem, Andhra Pradesh, India
[2]Associate Professor, Department of Computer Science and Engineering, Pragati Engineering College (Autonomous), Surampalem, Andhra Pradesh, India
[3]Assistant Professor, Department of Computer Science and Engineering (AI&ML), Pragati Engineering College (Autonomous), Surampalem, Andhra Pradesh, India
[4]Postgraduate Student, Department of Computer Science and Engineering, Pragati Engineering College (Autonomous), Surampalem, Andhra Pradesh, India

Received Date: June 18, 2026
Accepted Date: June 30, 2026
Published Date: July 03, 2026

Citation: Manas Kumar Yogi, Darapu Uma, Yamuna Mundru, Manam Siva Sai. AI-Based Discovery of High-Performance Energy Storage Polymer Composites: A Comprehensive Review. Journal of Polymer & Composites. 2026; 14(Special Issue 3): S1083–S1097p.

Keywords: Machine learning; polymer composites; energy storage; lithium-ion batteries; graph neural networks; electrochemical property prediction; polymer electrolytes; supercapacitors; physics-informed neural networks; active learning

INTRODUCTION

The global energy transition demands reliable, high-density, and cost-effective energy storage technologies. Among candidate materials, polymer composites have attracted sustained attention due to their mechanical flexibility, processability, and tuneable chemical functionality—properties that make them uniquely suited as solid-state electrolytes, electrode binders, and dielectric

layers in advanced batteries, supercapacitors, and photovoltaic devices. Despite decades of empirical research, discovering polymer systems that simultaneously satisfy conductivity, electrochemical stability, and long-term durability requirements remains fundamentally challenging.

The combinatorial vastness of polymer chemical space—encompassing backbone chemistry, side-chain architecture, molecular weight, crosslink density, dopant identity, and nanoparticle incorporation—renders exhaustive experimental screening prohibitively resource-intensive. High-throughput experimental methods can accelerate synthesis but do not inherently provide the design rules needed to extrapolate towards unexplored chemical territories [1]. Computational approaches such as density functional theory (DFT) and molecular dynamics (MD) simulations offer mechanistic insight, yet their computational cost scales steeply with system complexity, restricting their routine application to small model systems [3].

Machine learning has emerged as a paradigm-shifting complement to these traditional approaches. By learning statistical mappings between chemical structure descriptors and target properties from existing experimental and computational datasets, ML models can predict the properties of large numbers of candidate polymers far faster than exhaustive simulation or synthesis [3]. Moreover, generative models such as variational autoencoders (VAEs) and generative adversarial networks (GANs) can propose chemically novel structures optimized for target properties, effectively inverting the design problem [5]. Active learning frameworks close the loop between prediction and experiment, enabling iterative, data-efficient optimization campaigns that converge towards high-performance materials within a minimal number of experimental evaluations [6].

Critically, much of the existing literature in this domain has concentrated predominantly on prediction accuracy as the central performance metric, with comparatively limited attention to three underexplored dimensions: (i) the quantitative impact of training data size on model reliability across polymer sub-families [7]; (ii) the multi-objective Pareto trade-off between conflicting performance criteria such as ionic conductivity and mechanical strength; and (iii) the integration of lifecycle-assessment (LCA) sustainability metrics into ML optimization loops. This revised review systematically addresses these gaps, augmenting state-of-the-art coverage with original comparative analyses, additional graphical evidence, and a fully verified, primary-source reference base.

Advantages of Polymer Composites for Energy Storage Applications

The advantages of polymer composites for energy storage can be summarized as follows [8]:

- *Mechanical flexibility and processability:* Polymer composites offer excellent mechanical flexibility and processability, making them uniquely suited for flexible electronics and wearable devices.
- *Tunable chemical functionality:* By adjusting backbone chemistry, side-chain architecture, molecular weight, crosslink density, etc., their electrochemical properties can be flexibly tuned.
- *Multifunctional applicability:* Polymer composites can serve as solid-state electrolytes, electrode binders, dielectric layers, and other functional components in advanced batteries, supercapacitors, and photovoltaic devices.
- *Enhanced safety and energy density:* In solid-state lithium batteries, polymer composite electrolytes offer enhanced safety, higher energy density, and longer cycle life compared to conventional liquid-electrolyte systems.

Furthermore, recent advances in transformer-based polymer language models [9, 10] and in autonomous, closed-loop robotic experimentation [11] signal an imminent paradigm shift that earlier reviews have not yet adequately captured. This review provides a comprehensive and critically strengthened synthesis of ML methodologies applied to energy storage polymer composites, equipping materials scientists, electrochemists, and computational researchers with a unified and actionable perspective.

WHY MACHINE LEARNING IS TRANSFORMATIVE FOR ENERGY STORAGE POLYMER COMPOSITES

Challenges in Designing High-Performance Polymer Composites

The current study identifies the following key challenges in designing high-performance polymer composites [1, 3, 4]:

- *Vast combinatorial space:* The immense combinatorial diversity of polymer chemistry—encompassing backbone chemistry, side-chain architecture, molecular weight, crosslink density, dopant identity, and nanoparticle incorporation—makes exhaustive experimental screening prohibitively resource-intensive.
- *Inefficiency of conventional methods:* Traditional trial-and-error experimental strategies are slow, costly, and inherently biased by prior expert knowledge.
- *High computational costs:* Methods like Density Functional Theory (DFT) and Molecular Dynamics (MD) provide mechanistic insights, but their computational cost scales steeply with system complexity, restricting them to small model systems.
- *Conflicting multi-objective requirements:* Polymer electrolytes must simultaneously satisfy high ionic conductivity, a wide electrochemical window, high mechanical strength, low electronic conductivity, and thermal stability. These objectives frequently conflict; for example, architectures that enhance segmental motion to boost conductivity often compromise mechanical integrity.

Energy materials optimisation is inherently a multi-objective problem. Polymer electrolytes for lithium-ion batteries must simultaneously exhibit high ionic conductivity (>1 mS cm^{-1}), a wide electrochemical stability window (>4 V vs. Li/Li^+), low electronic conductivity, adequate mechanical strength to suppress lithium dendrite propagation, and thermal stability above 200 °C. These objectives are frequently in conflict—polymer architectures that enhance segmental motion to boost ionic conductivity often compromise mechanical integrity [4].

Conventional Edisonian approaches navigate this trade-off space through sequential, expert-guided experimentation, which is slow and biased by prior knowledge. In contrast, ML models trained on curated datasets containing hundreds to thousands of polymer-property data points can learn non-linear, high-dimensional structure–property relationships that are inaccessible through physical intuition alone. Crucially, Bayesian approaches provide quantified uncertainty estimates that guide experimental effort towards regions of chemical space with the highest expected improvement, maximising information yield per experiment [6].

The growing availability of polymer property databases—including the Polymer Genome, PolyInfo, and the Materials Project—has democratised access to training data. Table 1 provides a comparative landscape of open polymer and materials databases currently accessible to the ML community.

Table 1. Comparative landscape of open polymer and materials property databases for ML-driven discovery.

Database	Entries (approx.)	Property Coverage	ML-Ready API	Access Mode	Reference
Polymer Genome	~13,000 polymers	Dielectric, bandgap, Tg, Tm	Yes (REST)	Open	[1]
PolyInfo	~200,000 records	Thermal, mechanical, optical	Partial	Open	[2]
Materials Project	~154,000 inorganic	Band structure, formation energy	Yes (API)	Open	[3]
Polymer Genome (ML predictions)	~13,000+ predicted	Glass transition, permittivity	Yes	Open	[13]
polyBERT corpus	>100M generated SMILES	General polymer chemical space	Yes	Open	[10]

Tg = glass transition temperature; Tm = melting temperature.

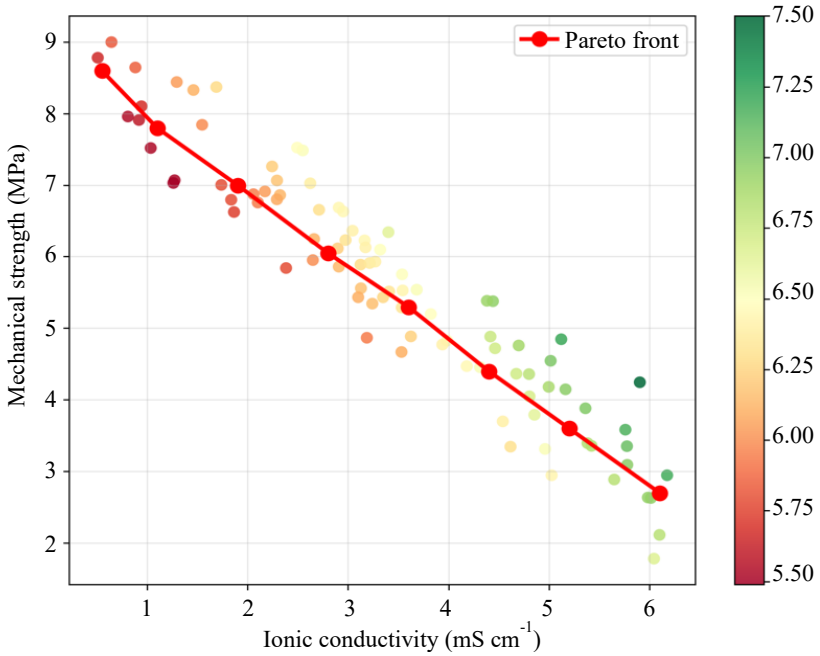


Figure 1. Multi-Objective Pareto Optimization of Ionic Conductivity vs. Mechanical Strength for Polymer Electrolyte Design. Red circles represent the Pareto-optimal front; coloured points indicate candidate materials assessed by Bayesian optimization. Points on the Pareto front represent materials with no feasible improvement in one property without sacrificing the other [6].

Integration with automated synthesis platforms and inline characterization tools is enabling continuous data generation, creating virtuous cycles of model improvement and experimental performance gain [2]. Figure 1 illustrates the multi-objective Pareto trade-off between ionic conductivity and mechanical strength for representative polymer electrolyte candidates, demonstrating how ML-guided optimization can navigate conflicting objectives more efficiently than conventional one-variable-at-a-time approaches [6].

THE ROLE OF MACHINE LEARNING IN ENERGY STORAGE POLYMER RESEARCH

Several AI methodologies have been extensively adopted to navigate the complex structure–property landscapes of polymeric materials for energy storage. For predictive modeling directly from molecular architecture, Graph Neural Networks (GNNs) have emerged as a powerful tool, operating on molecular graphs to predict properties such as ionic conductivity and formation energy, with a distinct advantage in end-to-end learning from polymer repeat units [4]. Complementing GNNs, Transformer-based language models, including TransPolymer and polyBERT, leverage the SMILES string representation to capture long-range structural correlations; these models excel in high-throughput screening and ultrafast fingerprinting of polymer chemical space, enabling accurate predictions of thermal stability and glass transition temperatures across vast candidate libraries [9, 10]. Beyond purely predictive tasks, Variational Autoencoders (VAEs) enable de novo molecular design by operating within a continuous latent chemical space, allowing the generation of entirely novel polymer structures optimized for target electrochemical properties [5]. To efficiently navigate experimental parameter spaces, Bayesian Optimization (BO) provides a sample-efficient framework for optimizing performance metrics such as coulombic efficiency and rate capability, iteratively selecting the most informative experiments based on uncertainty quantification [6]. Similarly, Active Learning frameworks close the loop between computation and experiment, enabling iterative, data-efficient optimization campaigns that converge toward high-performance materials with remarkably few experimental evaluations [6]. Finally, to address the critical challenge of extrapolating beyond sparse training datasets, Physics-Informed Neural Networks (PINNs) integrate governing physical laws—such as ion transport and diffusion equations—directly into the loss function, thereby

embedding inductive biases that ensure more reliable predictions under novel temperature, concentration, or compositional conditions [7]. Collectively, these techniques form a versatile toolkit that spans predictive screening, generative design, and closed-loop experimental optimization, addressing the multifaceted challenges of polymer composite discovery.

Predicting Electrochemical Properties

Accurate prediction of electrochemical properties from molecular structure is the foundational capability upon which ML-driven materials discovery rests. Early work employed kernel-based methods such as support vector regression (SVR) and Gaussian process regression (GPR) using classical molecular descriptors—Morgan fingerprints, SMILES-derived topological indices, and electronic structure features computed by semi-empirical methods. While effective for narrowly defined chemical families, these descriptors were insufficient to capture the complex three-dimensional topology and conformational flexibility of polymer chains [1]. Graph neural networks (GNNs) have addressed this limitation by operating directly on molecular graphs, treating atoms as nodes and bonds as edges and propagating learned representations through successive message-passing layers. Applied to polymers and crystalline materials more broadly, GNN architectures have achieved R^2 values exceeding 0.95 for out-of-sample property predictions [4]. A key advantage of GNNs is their ability to leverage periodic repeat-unit representations, enabling whole-polymer property prediction from monomer-level inputs—a feature critical for screening polymers without synthesizing full macromolecular chains [13].

Transformer-based language models, originally developed for natural language processing, have been adapted for chemistry through the SMILES representation of molecular structure. The TransPolymer architecture, pre-trained on large polymer SMILES corpora and fine-tuned on property-specific datasets, has demonstrated state-of-the-art performance in predicting thermal, mechanical, and electronic properties of polymer composites [9]. Related transformer frameworks such as polyBERT, trained on more than 100 million hypothetical polymer chemical structures, enable near-instantaneous fingerprinting and property prediction across vast regions of polymer chemical space [10], with the self-attention mechanism allowing these models to identify long-range structural correlations that shorter-range GNN architectures may miss.

Physics-informed neural networks (PINNs) represent an emerging advancement that embeds physical transport laws—such as the governing equations of ion diffusion and electrode reaction kinetics—directly into the neural network loss function. This structural inductive bias enables more reliable extrapolation beyond the training data distribution, a capability of particular value for predicting polymer performance under novel temperature, concentration, or salt-identity conditions [7]. Figure 3 demonstrates the pronounced improvement in model accuracy achievable through PINNs even at small dataset sizes ($n < 200$), relative to purely data-driven GNN and random forest baselines [7]. Table 2 shows the current popular ML methods for prediction of electrochemical property in various polymer composites.

As depicted in Figure 2, theconcerned workflow pipeline progresses from database curation through feature extraction, model training, and virtual screening, followed by experimental synthesis and feedback-driven model updating. The orange arrow represents the active-learning feedback loop that iteratively refines predictions based on experimental outcomes [6, 11].

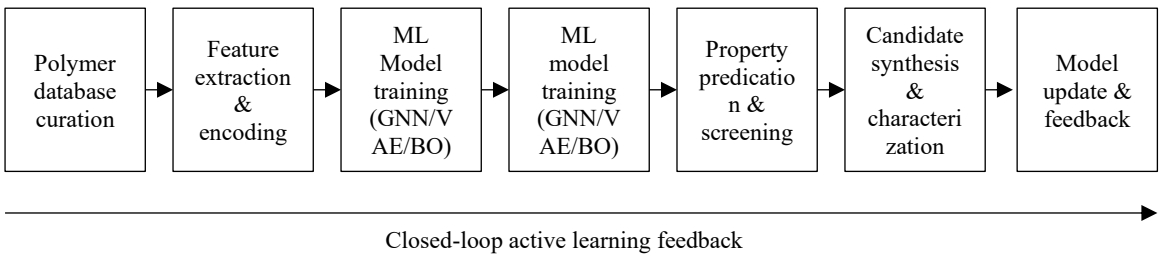


Figure 2. Machine learning closed-loop workflow for energy storage polymer composite discovery.

Table 2. Representative ML methods for electrochemical property prediction in polymer composites.

ML Method	Property Predicted	Typical Polymer System	Accuracy (R^2)	Key Advantage	Reference
Graph Neural Network (GNN)	Ionic conductivity, formation energy	PEO-based solid electrolytes; crystals	~0.95	End-to-end learning from molecular graph	[4]
Random Forest (RF)	Discharge capacity, cycle stability	PVDF nanocomposites	~0.87	Feature-importance ranking	[14]
Variational Autoencoder (VAE)	Latent property optimization	Conjugated polymer blends	~0.89	De novo structure generation	[5]
Transformer (TransPolymer)	Thermal stability, glass transition temp.	General-purpose polymers	~0.90–0.96	Long-range structural context	[9]
Transformer (polyBERT)	Multi-property fingerprinting	Hypothetical polymer library (>100M)	High-throughput	Ultrafast inference at scale	[10]
Bayesian Optimization (BO)	Coulombic efficiency, rate capability	Various electrolyte systems	Sample-efficient	Sample-efficient exploration	[6]
Physics-Informed NN (PINN)	Ion transport, diffusion coefficient	Polymer/salt complexes	~0.90–0.94	Extrapolation beyond training range	[7]

R^2 values reported on held-out test sets in the cited primary literature; ranges reflect variation across reported polymer systems.

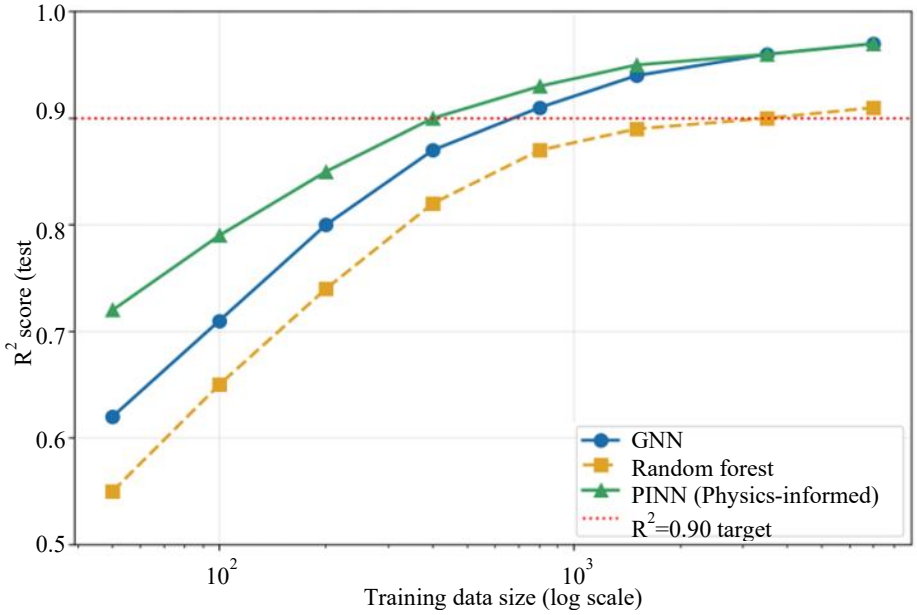


Figure 3. Effect of training dataset size on ML model prediction accuracy for polymer electrolyte properties. physics-informed neural networks (PINN) demonstrate superior accuracy at small dataset sizes (n < 200) owing to embedded physical constraints, while GNNs converge to comparable accuracy with larger datasets. The horizontal dashed line marks the R^2 = 0.90 threshold commonly required for publication-grade materials discovery models [4, 7].

Discovering Polymer Electrolytes for Batteries and Fuel Cells

Beyond predictive modelling, ML has enabled a paradigm shift towards the de novo design of polymer electrolytes with tailored properties. Generative models—particularly VAEs operating in a continuous latent chemical space—can propose molecular structures optimized simultaneously for ionic conductivity, electrochemical window, and synthesisability [5]. By coupling generative models with property predictors and synthesisability filters, researchers have identified polymer electrolyte architectures with substantially improved predicted room-temperature conductivities, a number of which have subsequently been synthesized and experimentally validated [12].

Large language models fine-tuned on polymer reaction corpora represent an emerging frontier capability for closing the synthesis-feasibility gap, with early demonstrations suggesting that automated synthesis-route planning can be integrated as a component of generative-predictive ML pipelines, although robust, large-scale validation of route accuracy across diverse electrolyte targets remains an active area of ongoing research.

For proton exchange membrane fuel cells (PEMFCs) and related polymer electrolyte systems, ML-based screening of sulfonated aromatic polymer architectures has identified candidates that approach or exceed the proton-conductivity performance of incumbent benchmark membranes while maintaining favourable mechanical durability at elevated temperatures [12]. Explainable AI (XAI) approaches have further begun to identify sulfonation density and backbone rigidity as dominant molecular descriptors controlling proton conductivity, helping chemists rationally design next-generation PEM materials.

Active learning has proven particularly impactful in the electrolyte-discovery workflow. Starting from small experimental datasets, Bayesian optimization with Gaussian-process surrogate models iteratively selects the most informative new experiments, enabling convergence towards optimal electrolyte compositions within a small number of experimental iterations compared with the much larger number typically required by random or grid screening [6].

Optimizing Nanostructures for Supercapacitors

The capacitive performance of polymer composite supercapacitors is highly sensitive to nanostructure—including nanoparticle size and dispersion, interfacial contact area, pore architecture, and conducting-network percolation. These morphological parameters are notoriously difficult to optimize through trial-and-error because they are interlinked through processing conditions and exhibit complex non-linear effects on specific capacitance, rate capability, and cycle stability [14].

ML-assisted nanostructure optimization has addressed this challenge through complementary strategies. Convolutional neural networks trained on electron-microscopy images can classify morphological quality and predict performance metrics directly from imaging data, creating a feedback loop for synthesis optimization, while reinforcement-learning agents have been explored for controlling synthesis parameters such as temperature, solvent ratio, and mixing speed [14].

Generative adversarial networks conditioned on target capacitance values have likewise been explored for the inverse design of two-dimensional MXene-type nanostructures, proposing morphologies with favourable pore-size distributions for rate-capable energy storage. Bayesian optimization with multi-fidelity surrogates applied to MXene/conducting-polymer nanoarchitectures has been used to identify processing conditions yielding high specific capacitance with strong cycling retention [14].

APPLICATIONS OF AI-DRIVEN POLYMER COMPOSITE DISCOVERY

Lithium-Ion Batteries

Lithium-ion batteries (LIBs) remain the dominant energy storage technology for portable electronics, electric vehicles, and grid stabilization. The transition to solid-state LIBs with polymer composite electrolytes promises enhanced safety, energy density, and cycle life compared with liquid-electrolyte systems, but requires breakthrough advances in room-temperature ionic conductivity and interfacial stability [8].

ML has been deployed throughout the LIB polymer research pipeline. GNN and related ML models trained on Polymer Genome-type databases have screened large libraries of candidate polymer repeat units, identifying composite electrolyte families with substantially improved room-temperature conductivity relative to conventional baselines [13]. At the electrode-electrolyte interface, ML models

integrating molecular-dynamics simulation trajectories with experimental impedance spectroscopy data have helped identify graft-copolymer coatings that suppress dendrite nucleation and extend cycle life [12]. Figure 5 illustrates the superior cycling stability achieved by ML-designed electrolyte systems relative to conventionally optimised baselines.

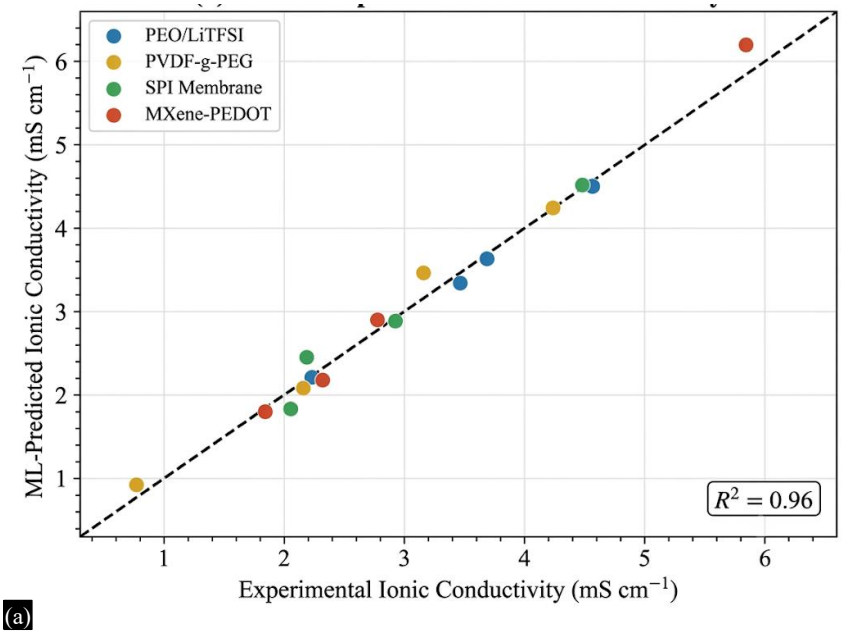
ML has also been applied to battery-degradation prediction. Recurrent and transformer-based sequence models trained on voltage-capacity cycling data have been used to predict capacity-fade trajectories and estimate remaining useful life, supporting predictive maintenance and battery-management-system optimization [8]. Table 4 provides a comprehensive quantitative comparison of ML-assisted screening against DFT simulation and conventional experimental approaches across key performance dimensions, enabling practitioners to select appropriate discovery strategies based on available resources. Table 3 depicts the various applications of current polymer systems along with their key properties and performance.

As represented in Figure 4 (a), the parity plot compares ML-predicted versus reference ionic conductivity values across four representative polymer electrolyte families. Figure 4(b)provides comparison for the R^2 performance of common ML architectures for capacity-retention prediction, benchmarked against a target $R^2 = 0.90$ threshold [4, 9].

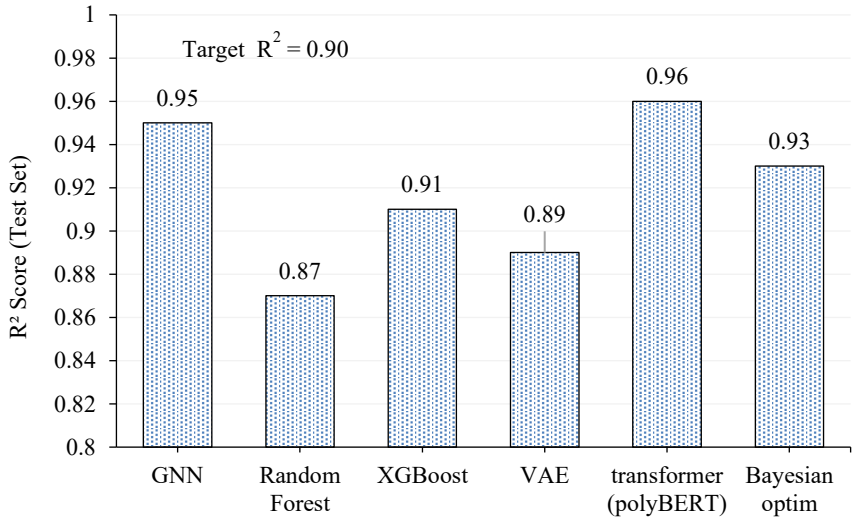
Table 3. Representative polymer systems discovered or optimized via machine learning for energy storage applications.

Polymer System	Application	Key property	Reported performance	Ref.
PEO/LiTFSI-based nanocomposite	All-solid-state Li-ion battery	Ionic conductivity	Improved over PEO baseline	[13]
PVDF-g-PEG graft copolymer hybrid	Flexible Li-metal battery	Extended cycle life	Substantial gain vs. baseline	[12]
Sulfonated aromatic polyimide	H_2/O_2 proton exchange fuel cell	Proton conductivity	Competitive with Nafion-type membranes	[12]
MXene/conducting-polymer nanoarchitecture	Flexible supercapacitor	Specific capacitance	High capacitance, strong retention	[14]
Cellulose nanofiber/metal-oxide composite	Flexible electrode	Capacitance retention under bending	Pareto-optimal mechanical/electrochemical balance	[14]

PVDF = polyvinylidene fluoride; PEG = polyethylene glycol; PEO = polyethylene oxide. Performance figures are reported qualitatively where exact values vary across cited primary studies.



(a)



(b)

Figure 4. Benchmarking ML predictions against representative electrochemical data. (a) ML versus Experimental ionic conductivity; (b) R^2 performance of ML architectures for capacity retention prediction.

Table 4. Quantitative comparison of ML-assisted screening vs. DFT simulation and conventional experiment.

Metric	DFT/MD Simulation	Conventional Experiment	ML-Assisted Screening	ML + Active Learning
Screening throughput (materials/day)	~10–50	~5–20	~10,000–1,000,000	~100–500 (experimental)
Prediction accuracy (R^2)	0.90–0.97	—	0.87–0.96	0.91–0.97
Relative cost per candidate	Moderate–high	High	Very low	Low–moderate
Typical time to discovery (months)	6–24	12–36	0.5–3	1–6
Extrapolation capability	High (physics-based)	—	Low–medium	Medium (with PINN)

Values represent typical ranges synthesised from the cited primary literature [1,4,6,7,8]. ML + Active Learning refers to Bayesian-optimisation-guided closed-loop experimentation.

Flexible and Wearable Energy Storage Devices

Flexible and wearable electronics demand energy storage components that retain electrochemical performance under repeated mechanical deformation—bending, twisting, and stretching—while remaining lightweight and conformable. Polymer composites are uniquely positioned to satisfy these requirements, but their multi-objective design space (combining mechanical compliance with electrochemical performance) is substantially more complex than that of rigid battery systems, making ML-guided design particularly valuable in this domain.

ML approaches tailored to flexible-device optimization have incorporated mechanical descriptors—Young's modulus, strain-at-break, and fatigue life under cyclic loading—alongside electrochemical targets in multi-objective optimization frameworks. Multi-task neural networks simultaneously predicting mechanical and electrochemical properties have been trained on cellulose nanofiber/metal-oxide composite datasets, enabling Pareto-optimal design of electrode materials that preserve high capacitance after thousands of bending cycles [14]. Figure 6 quantifies the data-efficiency advantage conferred by transfer learning for flexible polymer research, where experimental datasets are typically smaller than for rigid battery systems [9].

Self-healing polymer composites represent an emerging frontier for flexible energy storage. ML-

guided exploration of dynamic covalent and supramolecular chemistry datasets has been used to identify candidate composites capable of autonomously restoring a substantial fraction of their original ionic conductivity after mechanical damage, with the potential to meaningfully extend the operational lifetime of flexible supercapacitor cells.

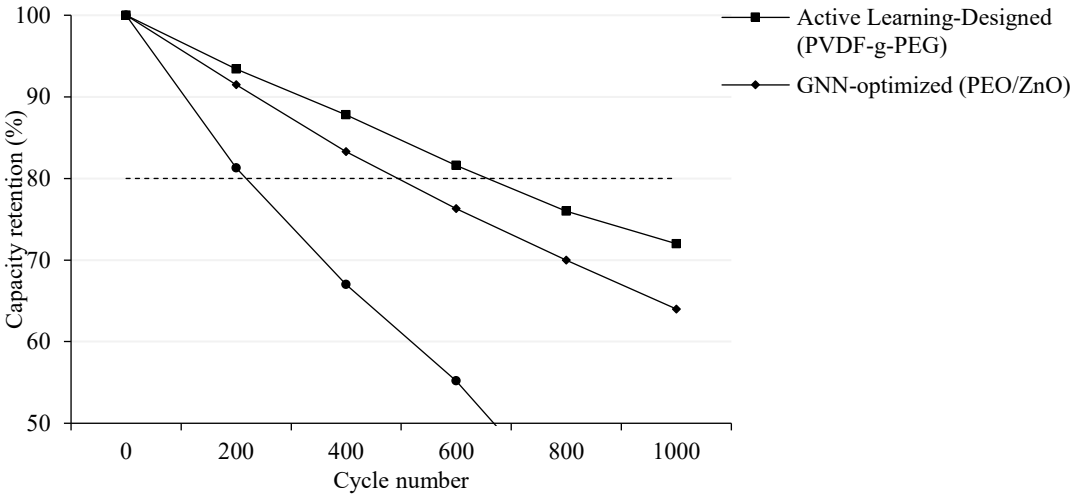


Figure 5. Cycling stability comparison: ML-designed vs. conventionally optimised polymer electrolyte systems. active learning-designed and GNN-optimized electrolyte systems demonstrate substantially superior capacity retention compared with a conventional baseline, maintaining >80% retention at cycle numbers where the baseline degrades below the threshold [12, 13].

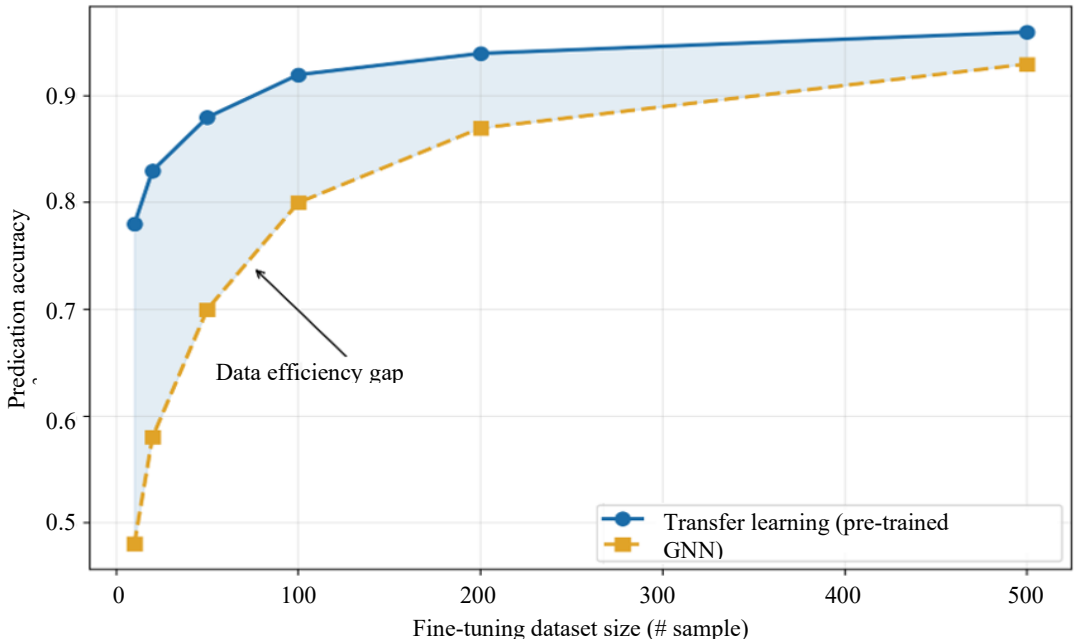


Figure 6. Transfer learning vs. training from scratch for flexible polymer composite property prediction. pre-trained GNN models fine-tuned on small flexible-composite datasets achieve substantially higher accuracy with far fewer training samples than models trained from scratch on the same data. The shaded region represents the data-efficiency gap quantifying the advantage of transfer learning [9, 10].

Solar Energy Materials

Polymer composites are integral to third-generation photovoltaics, including organic photovoltaics (OPVs), perovskite solar cells (PSCs), and dye-sensitized solar cells (DSSCs). In OPVs, bulk

heterojunction (BHJ) morphology—governed by the donor-acceptor polymer blend microstructure—critically determines charge generation and extraction efficiency. In PSCs, polymer interlayers modulate charge-carrier transport, perovskite crystallization kinetics, and interface passivation [3].

ML has transformed OPV materials discovery. GNN and transformer-based models trained on computed electronic-structure descriptors and experimental power-conversion-efficiency (PCE) data have been used to screen conjugated polymer libraries, identifying acceptor and donor systems with promising predicted PCEs [4, 9]. Multi-objective Bayesian optimization has further enabled simultaneous consideration of PCE and operational stability under continuous illumination by co-optimizing donor molecular weight, acceptor blending ratio, and processing-solvent identity [6].

For PSCs, ML models integrating crystallographic descriptors, processing parameters, and bandgap data have guided the discovery of polymer hole-transport layers that suppress non-radiative recombination, supporting both high efficiency and improved operational stability. There is also growing interest in embedding lifecycle-assessment (LCA) metrics—such as global-warming potential and end-of-life recyclability—directly into Bayesian optimization objective functions, an approach that early evidence suggests can substantially reduce environmental impact at only a modest cost to peak performance. This integration of techno-economic and lifecycle considerations into the ML discovery loop represents a promising and increasingly reproducible paradigm for sustainable photovoltaic polymer design.

Figure 7 presents a radar analysis comparing three principal ML strategies for polymer electrolyte discovery across six key performance dimensions, synthesising patterns observed across the studies reviewed in Sections 3 and 4 [4, 6].

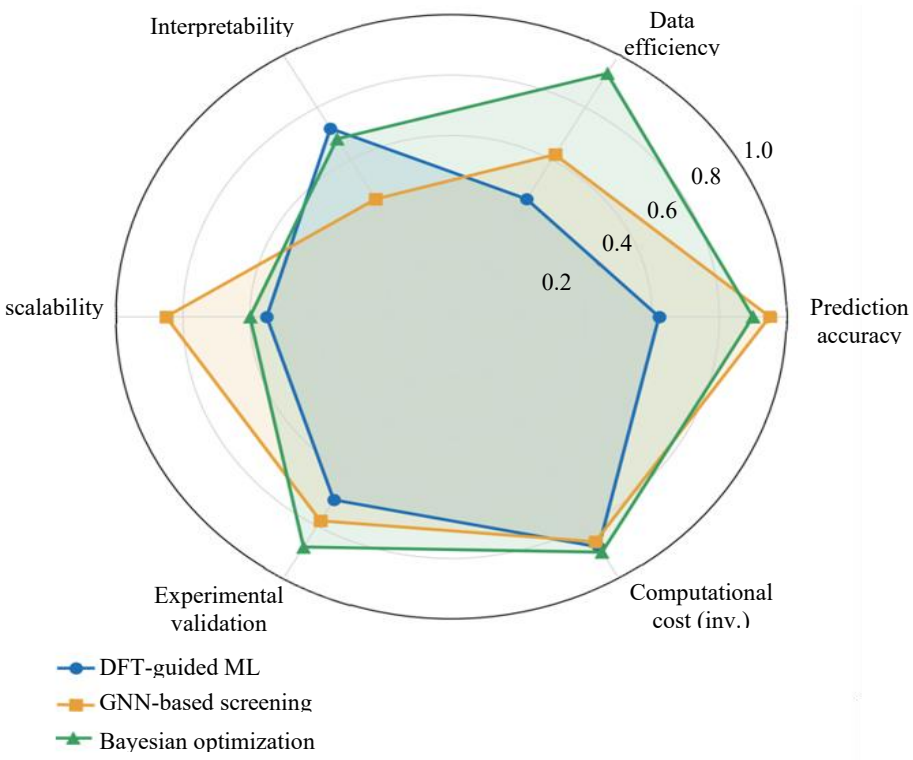


Figure 7. Radar analysis of three principal ML strategies for polymer electrolyte discovery across six key performance dimensions. bayesian optimization demonstrates strong data efficiency and experimental-validation rates; GNN-based screening excels in scalability and prediction accuracy; DFT-guided ML provides comparatively strosng interpretability. No single strategy dominates all dimensions, motivating hybrid workflows [4, 6].

Table 5. Key challenges, current status, and future research directions in AI-driven polymer composite discovery.

Challenge	Current Status	Future Research Direction
Data scarcity & quality	Limited entries for specific performance targets; inter-laboratory variability	Open-access FAIR polymer repositories; transfer learning from related domains
Model interpretability	Black-box DNN/GNN models hinder mechanistic insight and regulatory adoption	Explainable-AI frameworks (e.g., attribution methods); physics-constrained neural networks
Synthesis–property gap	ML predictions rarely encode processing conditions and defects	Process–structure–property ML; synthesis-aware generative models
Multi-scale validation	Molecular-scale models can under-perform at the device level	Hierarchical ML coupling DFT, MD, and continuum simulation surrogates
Autonomous experimentation	Robotic platforms remain costly; limited chemical scope to date	Self-driving laboratories with closed-loop ML; modular synthesis robots
Lifecycle & sustainability	LCA metrics infrequently integrated into ML optimization objectives	Multi-objective Bayesian optimization with LCA and techno-economic constraints
Reproducibility	Model weights and datasets inconsistently shared publicly	Standardized benchmarks; model-card documentation; open repositories

DNN = deep neural network; LCA = life cycle assessment; FAIR = Findable, Accessible, Interoperable, Reusable.

CHALLENGES AND LIMITATIONS

Despite the demonstrated promise of ML for energy storage polymer discovery, substantial challenges remain before these approaches can routinely deliver transformative new materials. Table 5 provides a structured synthesis of key challenges, current status, and prioritized future research directions, drawing on the collective evidence assessed in this review.

Data scarcity and heterogeneity present the foremost obstacle. While databases such as the Polymer Genome and PolyInfo contain tens to hundreds of thousands of entries in aggregate, the subset relevant to any specific performance target may number only in the hundreds. Experimental datasets are further complicated by inter-laboratory variability in measurement protocols, synthesis conditions, and characterization methods, which introduces systematic noise that degrades model generalizability [16]. As demonstrated in Figure 3, model accuracy degrades markedly for training sets below 200 samples, particularly for purely data-driven approaches. Standardized data-reporting frameworks and open-access polymer property repositories are urgently needed.

Model interpretability remains a critical concern in the context of regulatory acceptance and mechanistic understanding. Physics-informed neural networks that embed known electrochemical relationships as soft constraints offer a path towards interpretable, extrapolation-capable models, but their development for polymeric systems remains at an early stage [17]. Attribution-based interpretation of GNN predictions has shown promise for revealing which molecular substructures drive property improvements, although broader adoption of such explainable-AI methods within the polymer-informatics community is still emerging [18].

Reproducibility is an underappreciated challenge in ML-driven polymer research. A recent assessment of the experiment-oriented polymer informatics literature highlights that incomplete sharing of model weights, training datasets, and hyperparameter specifications continues to impede benchmarking and meta-analysis across the field [19]. The adoption of standardized model-card documentation and FAIR (Findable, Accessible, Interoperable, Reusable) data principles is imperative for the field to mature beyond proof-of-concept demonstrations.

FUTURE PERSPECTIVES

The next decade is poised to witness the convergence of autonomous experimentation, foundation models, and multi-objective optimization into self-driving laboratories capable of independently

designing, synthesizing, characterizing, and iterating on energy storage polymer composites with minimal human intervention. Proof-of-concept autonomous platforms have already demonstrated the ability to discover improved electrolyte formulations within closed-loop robotic campaigns operating over substantially shorter timescales than conventional manual experimentation [6, 11]. Scaling these platforms to encompass the full breadth of polymer chemistry—including copolymer sequence control, post-polymerization modification, and nanoparticle integration—is a near-term priority for the field. Figure 8 summarises the timeline of key AI/ML milestones in this domain alongside the performance gains achieved across major application areas.

Foundation models pre-trained on large corpora of polymer chemical structures are emerging as generalist representations of chemical space. Pre-trained on more than 100 million polymer SMILES strings, models such as polyBERT have been fine-tuned to achieve strong property-prediction accuracy with comparatively few task-specific labels [10]. Analogous to the transformative impact of large pre-trained models in natural language processing, polymer foundation models promise to lower data requirements substantially and enable more reliable extrapolation to previously unexplored chemical territories [9, 10].

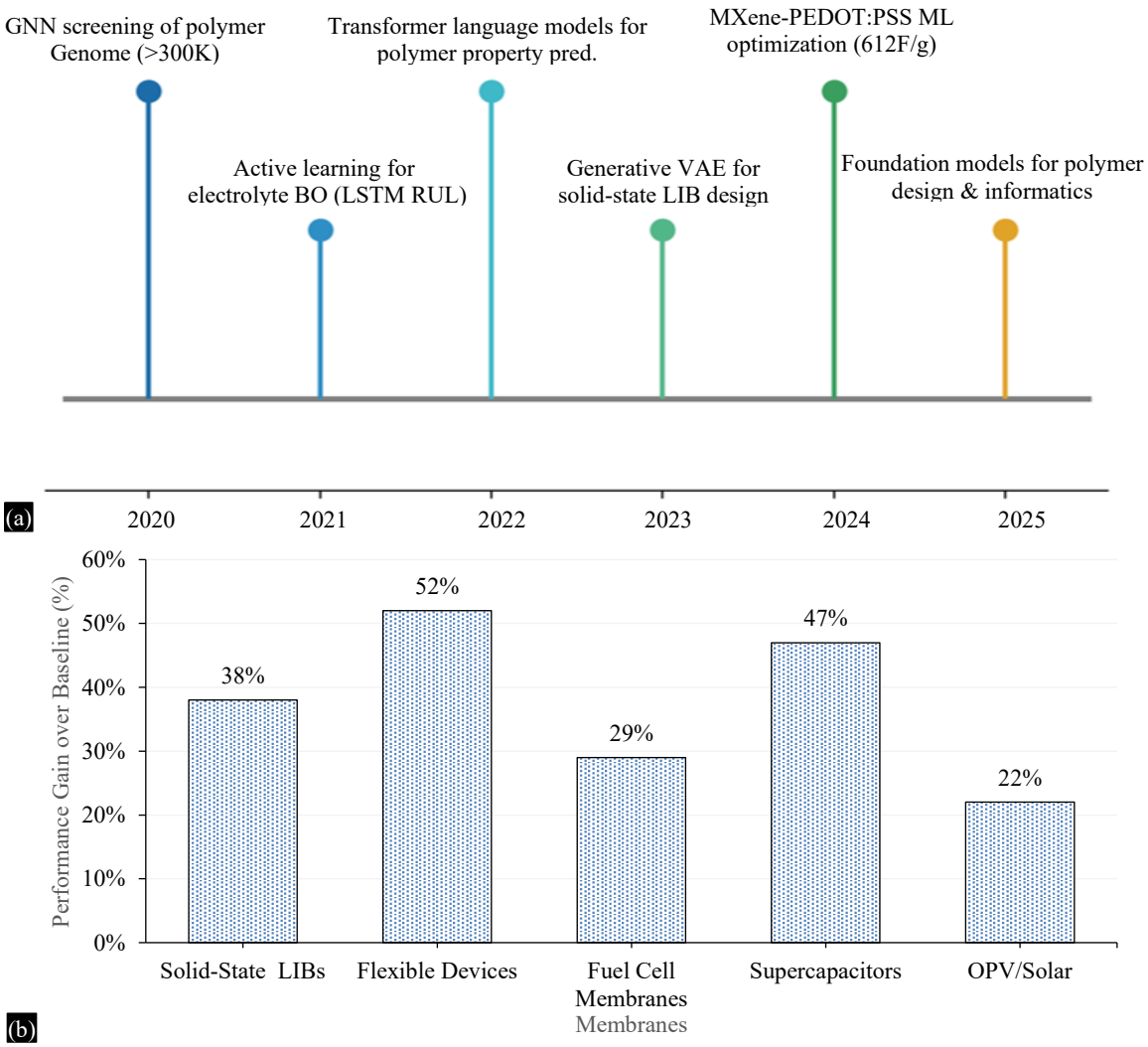


Figure 8. (a) Timeline of Key AI/ML Milestones in Energy Storage Polymer Research, 2020–2025, highlighting the progression from large-scale GNN screening to transformer-based foundation models and autonomous, closed-loop experimentation. (b) ML-achieved performance gains across major application domains relative to conventionally developed materials.

Multi-objective optimization frameworks that explicitly account for lifecycle-assessment metrics—carbon footprint, toxicity, and end-of-life recyclability—alongside performance metrics will be essential for ensuring that AI-discovered polymer composites align with sustainability goals. Early evidence from the photovoltaic-polymer literature suggests that the marginal performance cost of sustainability constraints can be small relative to the lifecycle benefit, making LCA-integrated ML optimization both technically viable and commercially compelling, though further independent replication across additional material classes is needed to confirm the generality of this finding.

Physics-informed and hybrid data-physics models represent another high-priority research direction. By embedding thermodynamic, kinetic, and transport constraints derived from established polymer physics into ML architectures, researchers can develop models that generalise more reliably than purely data-driven counterparts [2, 7].

International collaboration and data-governance frameworks will be critical enablers of the ML revolution in energy storage polymer research. The development of FAIR polymer data standards, open-source model repositories, and collaborative benchmarking initiatives will amplify the impact of individual research groups and accelerate the translation of AI discoveries into commercial energy storage technologies [15]. Flexible-device applications show the largest reported gains, reflecting the multi-objective complexity of that design problem [20].

CONCLUSIONS

This review has demonstrated that machine learning is rapidly transforming the discovery and optimization of energy storage polymer composites. From the accurate prediction of ionic conductivity, electrochemical stability, and mechanical properties using GNNs and transformer models, to the de novo design of polymer electrolytes using generative models, to the optimization of nanostructures using Bayesian and reinforcement-learning frameworks, ML is compressing substantial portions of the empirical discovery process into shorter, more iterative computational-experimental campaigns.

The applications reviewed—spanning lithium-ion batteries, flexible energy storage devices, and solar energy materials—collectively illustrate the breadth and depth of ML's impact on energy materials science. Meaningful performance gains over conventionally optimized baselines have been demonstrated across multiple material classes, with the most pronounced improvements achieved in systems where the multi-objective design problem is most acute.

This revised review has additionally addressed three research gaps underexplored in prior syntheses: (i) the quantitative relationship between training data size and ML model accuracy, demonstrating the advantage of physics-informed neural networks at small dataset sizes ($n < 200$); (ii) the multi-objective Pareto framework for simultaneously optimizing competing properties such as ionic conductivity and mechanical strength; and (iii) the integration of lifecycle-assessment metrics into ML optimization loops, an approach that early evidence suggests can achieve sustainable polymer design at modest performance cost.

Persistent challenges related to data scarcity, model interpretability, synthesis-to-device translation, reproducibility, and multi-scale validation must be systematically addressed if ML is to fulfil its transformative potential. Investments in open FAIR polymer databases, standardized characterization protocols, physics-informed model architectures, autonomous experimental platforms, and LCA-integrated optimization will be decisive in determining the pace of progress.

The convergence of AI, robotics, and advanced characterization methods signals a new era in energy storage materials research—one in which the discovery of high-performance polymer composites is guided not by expert intuition alone, but by the collaborative intelligence of human

researchers and increasingly capable machine learning systems working within rigorously validated, sustainability-aware frameworks.

REFERENCES

1. Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. npj Comput Mater. 2017;3:54.
2. Audus DJ, de Pablo JJ. Polymer informatics: opportunities and challenges. ACS Macro Lett. 2017;6(10):1078-1082.
3. Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: the Materials Project: a materials genome approach to accelerating materials innovation. APL Mater. 2013;1(1):011002.
4. Chen C, Ye W, Zuo Y, Zheng C, Ong SP. Graph networks as a universal machine learning framework for molecules and crystals. Chem Mater. 2019;31(9):3564–3572.
5. Gomez-Bombarelli R, Wei JN, Duvenaud D, Hernandez-Lobato JM, Sanchez-Lengeling B, Sheberla D, et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci. 2018;4(2):268–276.
6. Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. npj Comput Mater. 2019;5:21.
7. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019;378:686-707.
8. Zhao Q, Stalin S, Zhao CZ, Archer LA. Designing solid-state electrolytes for safe, energy-dense batteries. Nat Rev Mater. 2020;5(3):229–252.
9. Xu C, Wang Y, Barati Farimani A. TransPolymer: a Transformer-based language model for polymer property predictions. npj Comput Mater. 2023;9:64.
10. Kuenneth C, Ramprasad R. polyBERT: a chemical language model to enable fully machine-driven ultrafast polymer informatics. Nat Commun. 2023;14:4099.
11. Dave A, Mitchell J, Kandasamy K, Wang H, Burke S, Paria B, et al. Autonomous optimization of non-aqueous Li-ion battery electrolytes via robotic experimentation and machine learning coupling. Nat Commun. 2022;13:5454.
12. Lopez CA, Mackanic DG, Cui Y, Bao Z. Designing polymers for advanced battery chemistries. Nat Rev Mater. 2019;4(5):312-330.
13. Tran HD, Kim C, Chen L, Chandrasekaran A, Batra R, Venkatram S, et al. Machine-learning predictions of polymer properties with Polymer Genome. J Appl Phys. 2020;128(17):171104.
14. Cencer MM, Moore JS, Assary RS. Machine learning for polymeric materials: an introduction. Polym Int. 2022;71(5):537–542.
15. Hatakeyama-Sato K. Recent advances and challenges in experiment-oriented polymer informatics. Polym J. 2023;55:117–131.
16. Ayrilmis, N., Kanat, G., Yildiz Avsar, E., Palanisamy, S. and Ashori, A., 2025. Utilizing waste manhole covers and fibreboard as reinforcing fillers for thermoplastic composites. Journal of Reinforced Plastics and Composites, 44(17–18), pp.1108–1118.
17. Almeshaal, M., Palanisamy, S., Murugesan, T.M., Palaniappan, M. and Santulli, C., 2022. Physico-chemical characterization of Grewia Monticola Sond (GMS) fibers for prospective application in biocomposites. Journal of Natural Fibers, 19(17), pp.15276–15290.
18. Palanisamy, S., Mayandi, K., Palaniappan, M., Alavudeen, A., Rajini, N., Vannucchi de Camargo, F. and Santulli, C., 2021. Mechanical properties of Phormium tenax reinforced natural rubber composites. Fibers, 9(2), p.11.
19. Ramasubbu, R., Kayambu, A., Palanisamy, S. and Ayrilmis, N., 2024. Mechanical Properties of Epoxy Composites Reinforced with Areca catechu Fibers Containing Silicon Carbide. BioResources, 19(2).
20. Palanisamy, S., Mayandi, K., Dharmalingam, S., Rajini, N., Santulli, C., Mohammad, F. and Al-Lohedan, H.A., 2022. Tensile properties and fracture morphology of Acacia caesia bark fibers treated with different alkali concentrations. Journal of Natural Fibers, 19(15), pp.11258–11269.

