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

Research JoPC

Generative AI for Designing Sustainable Polymer Composites for Renewable Energy Applications

Joshila Grace L.K.[1,*], Kapil S. Banker[2], Harshit P. Bhavsar[3], Rohini Goel[4], Mit C. Patel[5], Bharatkumar D. Prajapati[6]

Abstract

Sustainable polymer composites are increasingly required for renewable energy devices, yet conventional trial-and-error formulation cannot efficiently balance performance, processability, recyclability, and environmental constraints. This study proposes a generative artificial intelligence framework for designing polymer composites for photovoltaic encapsulation, dielectric energy storage, polymer electrolytes, and thermal-management systems. Public polymer-property and composite datasets were curated from open databases and published supplementary records. Chemical descriptors, molecular fingerprints, polymer embeddings, processing variables, and sustainability indicators were used to train a masked multitask property predictor. Conditional generative models produced new polymer candidates, which were screened through chemical validity, novelty, synthesizability proxy, uncertainty estimation, and physics-guided composite rules. The multitask model achieved reliable prediction across thermal, mechanical, dielectric, ionic, and sustainability-related targets, with test-set R^2 values ranging from 0.79 to 0.92. From 20,000 generated candidates, 92.10% were chemically valid and 81.64% were novel. Final ranking identified promising bio-based, recyclable, and energy-functional composite classes. The proposed framework provides a reproducible route for early-stage sustainable polymer composite discovery and supports future experimental validation under real renewable energy operating and aging conditions before device scale-up.

Keywords: Generative artificial intelligence, sustainable polymers, polymer composites, renewable energy materials, polymer informatics, recyclable polymers, dielectric materials, polymer electrolytes

***Author for Correspondence**
Joshila Grace L.K.

[1]Professor, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India
[2,6]Assistant Professor, Department of Mechanical Engineering, Government Engineering College, Palanpur, Gujarat, India
[3]Associate Professor, Department of Mechanical Engineering, Swarrnim Institute of Technology, Swarrnim Startup & Innovation University, Gujarat, India
[4]Associate Professor, Department of Computer Science & Engineering, MM Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, Haryana, India
[5]Associate Professor, Department of Mechanical Engineering, Silver Oak University, Ahmedabad, Gujarat, India

Received Date: June 27, 2026
Accepted Date: July 06, 2026
Published Date: July 16, 2026

Citation: Joshila Grace L.K., Kapil S. Banker, Harshit P. Bhavsar, Rohini Goel, Mit C. Patel, Bharatkumar D. Prajapati. Generative AI for Designing Sustainable Polymer Composites for Renewable Energy Applications. Journal of Polymer & Composites. 2026; 14(Special Issue 3): S1232–S1257p.

INTRODUCTION

Background and Scientific Motivation

As the shift to renewable energy technologies has accelerated, the need for polymeric materials that are functional, processable and sustainable throughout their material life cycle has grown. In this context, polymer composites are in an important position as the polymer matrix can be chemically modified, mechanically reinforced and functionally modified by fillers, fibers, nanophases and/or interfacial engineering. The transition from petroleum-based plastics to sustainable polymer science has already progressed beyond just replacing the petroleum-based plastics, and now encompasses renewable monomers, recyclable thermosets, degradable architectures, low-carbon processing, and circular end-of-life pathways [1].

This change is particularly relevant for energy use applications as the material needs to meet the challenging service requirements and deliver the decarbonization targets. The real challenge is, therefore, not only to design a "green" polymer, but to design a polymer composite that is both sustainable and stable, has good performance, is manufacturable and is cost effective, all within the same formulation space. The use of polymer composites in renewable energy systems such as photovoltaic modules, wind turbine structures, fuel-cell membranes, supercapacitors, dielectric capacitors, solid polymer electrolytes, and protective coatings is gaining more and more importance. They are low density, corrosion resistant, tunable in their dielectric response, mechanically flexible, and can be fabricated at a scale that is compatible with the energy components [2]. For example, in photovoltaics, polymers are employed as encapsulants, flexible substrates, interfacial modifiers, charge-transport layers and protective coatings; in these applications, the resistance to moisture, to UV radiation, the thermal aging behavior and the adhesion to inorganic layers are important factors that affect the lifetime of the modules [3]. Likewise, polymer-based dielectric and nanocomposite systems are extensively studied for energy storage applications in capacitors due to their tunable breakdown strength, tunable dielectric constant, tunable charge–discharge efficiency and tunable interfacial polarization [4].

Problem Statement and Research Gap

With this wide range of potential, however, it is still a challenging materials-engineering challenge to design polymers for composites that are sustainable. The conventional experimental development relies heavily on trial and error formulation techniques, and optimization of polymer chemistry, filler loading, particle morphology, dispersion quality, crystallinity, glass-transition temperature, interphase thickness, processing temperature and curing conditions are all interdependent. A slight modification of the chemistry of the filler or its surface treatment can lead to an increase in ionic conductivity, but decrease in mechanical toughness; using a bio-based matrix can lead to a decrease in environmental impact, but increase in moisture sensitivity; a highly polar repeat unit can lead to an increase in the dielectric constant, but decrease in breakdown strength. These trade-offs are not linear and they can arise from multiscale interactions between the molecular structure, mesoscale morphology, interfacial adhesion and device level performance [5].

The sustainability aspect is an additional challenge. While renewable energy devices need materials that can last for a long time, there are many sustainable polymers that are still lacking in the property databases, uncertain in terms of recyclability when in composite form, lacking in thermal aging data, and the lack of standardization of life-cycle indicators. Energy-storage polymer composites, too, require high energy density and low dielectric loss, and do not allow for premature electrical treeing, thermal runaway or defect formation due to the presence of fillers [6]. Previous advances in polymer dielectrics and high-energy-density materials have demonstrated that molecular polarity, free volume, chain rigidity and nanofiller interfaces are key factors that determine the final energy-storage response [7]. However, the design rules that are available are still disorganized. While the majority of studies optimize a single property class, the optimization of electrical, thermal, mechanical, environmental and processing constraints are required for renewable energy applications.

Generative AI-Guided Design as the Proposed Direction

Generative AI is a new way to solve this combinatorial design problem. Instead of screening just known polymers or manually changing the recipe of composites, generative models can learn structure–property–processing relationships and suggest novel polymer repeat units, copolymer sequences, functional groups, filler combinations or composite architectures for a set of target conditions. Polymer informatics has already demonstrated that curated polymer databases, molecular fingerprints, graph representations, and surrogate learning models can accelerate property prediction and candidate screening [8]. Data-driven polymer design further enables researchers to connect chemical descriptors with glass transition temperature, dielectric constant, bandgap, thermal conductivity, modulus, solubility, and degradation behavior before expensive synthesis or fabrication begins [9].

Machine learning studies on polymer dielectrics have shown that predictive models can reduce the cost of exploring large chemical spaces and can support rational selection of candidates for energy applications [10]. Similar progress has been reported for frequency-dependent dielectric prediction, where polymer fingerprints and regression models were used to estimate dielectric behavior across wide operating regimes [11]. More recent graph neural network approaches improve the representation of polymer repeat units and support multi-property learning, which is valuable for composite design because renewable energy materials rarely depend on one isolated descriptor [12]. These advances indicate that a generative-AI framework, when coupled with physics-aware screening, can move polymer-composite research from forward prediction toward inverse design.

Novelty and Contribution of the Present Study

The novelty of the present work lies in positioning generative AI as a materials-design engine for sustainable polymer composites in renewable energy applications. Instead of treating polymer selection, filler selection, and sustainability assessment as separate tasks, the proposed study frames them as an integrated inverse-design problem. Chemical language models such as polyBERT have shown that polymer structures can be interpreted through sequence-based representations, making it possible to learn polymer chemistry in a manner similar to linguistic patterns [13]. Deep generative models, including variational auto encoders, recurrent models, reinforcement-guided generators, and graph-based generators, have also been benchmarked for de novo polymer design, revealing both their promise and their limitations in validity, novelty, and property targeting [14].

This paper therefore proposes a generative AI-assisted design concept in which sustainable polymer matrices, recyclable thermoset networks, bio-based monomers, functional nanofillers, and energy-specific performance targets are jointly considered. A recent work on the inverse design of recyclable polymers using AI has demonstrated the potential of using machine learning to propose experimentally relevant sustainable polymer chemistries [15]. This trend is further supported by foundation chemical language models, which can be used to generate specific polymers based on their desired properties [16]. Agentic polymer-design systems also suggest that natural-language interaction, property prediction and structure modification can be integrated in closed-loop polymer discovery workflows in the early stages of polymer design [17]. The present study seeks to establish a sustainable polymer-composite design pathway building on these developments, which involves generative candidate creation, property prediction, screening of the interface performance and sustainability-aware ranking. This is expected to decrease the amount of experiments, increase formulation efficiency and aid in the discovery of renewable energy polymer composites that have enhanced durability, processability and environmental compatibility [18].

LITERATURE REVIEW

Generative AI in Materials and Polymer Design

The recent advances of generative artificial intelligence have revolutionized the search, screening and optimization of materials. The previous computational materials design was primarily based on the density functional theory, molecular dynamics, or high throughput screening based on rules. These approaches are still useful, but are inefficient if the chemical design space is large. There is now another path, with generative models. They can identify patterns in already known structures of materials, and suggest new candidates that have specific properties. Choudhary presented AtomGPT, a generative pretrained transformer for forward and inverse materials design, demonstrating the ability of transformer-based models to bridge the gap between atomistic structure and prediction and candidate generation of properties [19]. This book is not only applicable to polymers, but also provides a valuable basis for polymer composite research as polymer design is also an inverse process from desired performance to feasible molecular or structural forms.

Nanocomposite design has been given significant attention for energy-related applications since the interface between the matrix and the filler is frequently the determinant of the thermal, electrical and mechanical properties. Jiang et al. reviewed nanocomposite systems for energy applications, and

highlighted the importance of the morphology, dispersion, interfacial bonding and phase compatibility of the fillers on the device level [20]. This is very relevant for the sustainable polymer composites since the materials used for renewable energy are seldom single-phase systems. They are typically hybrid structures with the polymer matrix, reinforcement and functional additive all playing a role.

AI-Assisted Composite Microstructure and Architecture Design

One of the challenges in the development of polymer-composites is the lack of ability to relate the microstructural features to the macroscopic response. This has begun to be tackled by deep generative models. Lee et al. suggested a diffusion-based generative approach to design multifunctional composite microstructures, using the learned microstructure representation to guide the property-oriented composite design [21]. This kind of approach is beneficial as many composite properties are not only dependent on the chemical composition of the polymer but also on the spatial arrangement, local continuity and interfacial topology.

The replacement of polymers has also come to the fore as a major theme of sustainability. To find all-natural substitutes for plastic, Chen et al. applied machine intelligence to help identify materials that have functional properties and are not harmful to the environment [22]. Kuenneth et al. went in a different direction to create a multitask deep neural network for the prediction of multiple polymer properties at once, instead of a single property [23, 24]. For polymer composites used in renewable-energy devices, this systems-level view is useful. Materials are no longer passive components. They interact with sensors, operating environments, control systems, and device-level monitoring platforms.

Composite Materials, Thermoelectrics, and Energy-Focused Discovery

Composite material discovery increasingly depends on hierarchical design. Park et al. introduced a hierarchical generative network for accelerated composite material discovery, where multitask learning helped connect material features with target performance [25]. This is important for polymer composites because formulation variables exist at different levels: monomer chemistry, chain architecture, filler chemistry, curing route, laminate sequence, and final device geometry.

AI-guided thermoelectric materials design also provides useful lessons for polymer-based renewable-energy systems. Han et al. reviewed AI methods for thermoelectric material discovery and highlighted the role of data-driven models in identifying complex property relationships [26]. Polymer composites for energy harvesting, dielectric storage, and thermal management face a similar problem. A high-performing material must show favorable transport, mechanical integrity, thermal stability, and long-term reliability at the same time.

Thermosetting resins are particularly important in structural renewable-energy systems, including wind turbine blades, electrical insulation, protective coatings, and composite housings. Zhao et al. employed machine learning and high throughput screening to design multicomponent thermosetting resins, which showed that AI can be used to cut the number of experiments needed to search in a very complex formulation space [27]. This is important because thermosets continue to be challenging to recycle, repair or reform after curing, making them sustainable thermosets. Improved design tools can be used to find resin systems that are tougher, less burdensome to process and may even be circular.

Renewable-Energy Systems and AI-Enabled Design Context

Evaluating polymer composites at the system scale as well as material scale is important for renewable-energy systems. Kishor et al. made a comparison between static and single axis solar tracking systems and demonstrated that the energy efficiency is not only dependent on the ideal material or device assumptions, but also on practical operating conditions [28]. The concept is applicable to polymer composites as coatings, encapsulants, substrates and structural polymers are required to function in the real solar, humidity, temperature cycling, dust and mechanical stress environments.

AI is also playing a significant role in battery technology, where it is being used to identify new materials for use in batteries. Li et al. talked about the multiscale AI innovation of batteries from material discovery to smart manufacturing [29]. This includes polymer electrolytes, separators, binders and encapsulation materials. They need to be designed in such a way that they are optimized for ionic transport, electrochemical stability, interfacial compatibility and mechanical robustness.

The discovery of polymers that are sustainable has also been a dedicated research area. Huo et al. summarized the machine learning methods for sustainable polymers and pointed out that future advances rely on the availability of more data, prediction of polymer properties, modeling of polymer degradation, and life-cycle-considered design [30]. Long et al. also reviewed the latest applications of AI in polymer materials, such as property prediction, structure design, and optimization of the processing [31]. The studies demonstrate that polymer informatics is progressing from the prediction of a single property to materials intelligence.

Research Gap and Need for the Present Study

Several works have linked AI with energy monitoring and renewable-energy operation. Gantla et al. presented a smart monitoring framework for sustainable solar energy harvesting using IoT and machine learning, showing the importance of data-driven feedback in energy systems [32]. Genetic algorithms and AI-based optimization have also been applied to product design, where design variables can be searched under multiple constraints [33]. In transport-energy systems, generative and explainable AI have been explored for electric-vehicle energy applications, indicating the wider relevance of interpretable AI in sustainable energy design [34]. Shin et al. used multimodal deep learning for PEMFC performance prediction, further confirming that renewable-energy materials and devices benefit from data-rich, model-assisted optimization [35].

As summarized in Table 1, existing studies have advanced AI-assisted polymer discovery, nanocomposite modeling, and sustainable resin design; however, most of them address these areas separately rather than as an integrated renewable-energy composite design problem. This gap supports the need for the present generative AI-guided framework, which jointly considers polymer chemistry, filler–matrix architecture, processability, recyclability, and energy-application performance.

However, the literature still shows a clear gap. Many studies focus on either polymer property prediction, nanocomposite modeling, renewable-energy device monitoring, or sustainable material selection. Few studies combine these ideas into a generative AI framework specifically for sustainable polymer composites in renewable-energy applications. A unified approach is therefore needed, where polymer chemistry, composite architecture, energy-performance requirements, processability, and sustainability indicators are considered together. This gap defines the motivation of the present work.

METHODOLOGY

Experimental Design and Workflow

The present study was designed as a public-data-based computational experiment for discovering sustainable polymer composites suitable for renewable energy applications. The methodology combined polymer informatics, generative molecular design, composite property screening, sustainability scoring, and multi-objective ranking. The workflow was not developed as a purely conceptual framework. Instead, it used publicly available polymer and materials data, reproducible preprocessing rules, fixed train–validation–test splits, and quantitative evaluation metrics.

The overall experimental pipeline is shown in Figure 1, where the design process begins with polymer-property data acquisition and ends with a ranked set of candidate polymer composites for photovoltaic, battery, supercapacitor, fuel-cell, dielectric, and thermal-management applications. Similar data-driven materials design principles have been used in polymer discovery, polymer electrolyte generation, composite modeling, and sustainable polymer screening [3, 6, 8, 30].

Table 1. Literature gap analysis and gap-to-solution mapping for generative AI-enabled sustainable polymer composite design.

S. N.	Author(s)/year/ ref.	Title/focus area	Methodology/tools used	Key findings	Limitations/Gaps identified	Relevance to current study
1	Wilson et al., 2023 [3]	AI discovery of sustainable polymers	PolyID platform, polymer informatics, property screening	Demonstrated AI-guided discovery of performance-advantaged sustainable polymers.	Focused mainly on polymer candidates; composite architecture and renewable-energy device constraints were limited.	Supports data-driven polymer selection as the first layer of the proposed framework.
2	Sharma et al., 2022 [8]	ML-assisted polymer composite design	Machine learning models for modeling, analysis, and design	Showed that ML can reduce trial-and-error in polymer composite development.	Limited integration of generative AI, sustainability metrics, and energy-specific performance targets.	Establishes the need for an AI-assisted composite design workflow.
3	Yang et al., 2024 [6]	De novo polymer electrolyte design	GPT-based and diffusion-based generative models	Generated polymer electrolyte candidates with targeted properties.	Application scope was mainly electrolyte chemistry; broader composite sustainability and filler effects were not central.	Provides a basis for generative polymer design under property constraints.
4	Zheng et al.,	Recyclable vitrimeric	AI-guided inverse design and	Identified recyclable vitrimeric polymers with improved	Did not fully address renewable-energy composite structures or	Supports the recyclable-polymer
	2025 [11]	Polymer discovery	experimental validation	Sustainability potential.	multifunctional device-level requirements.	dimension of the present study.
5	Jiang and Huang, 2025 [20]	Nanocomposites for energy applications	Review of nanofiller design, interfaces, and energy functions	Highlighted the role of filler dispersion, interphase behavior, and nanostructure in energy materials.	Mostly review-based; no unified AI-driven inverse-design route was proposed.	Justifies including filler–matrix interaction and interfacial screening.
6	Lee et al., 2024 [21]	Microstructure design of multifunctional composites	Diffusion-based deep generative model	Generated composite microstructures linked with multifunctional performance.	Microstructure generation was emphasized, while polymer chemistry and sustainability scoring were less developed.	Supports the microstructure-generation module of the proposed approach.
7	Kuenneth et al., 2022 [23]	Bioplastic design using multitask learning	Multitask deep neural networks	Predicted multiple bioplastic properties simultaneously.	Limited focus on renewable-energy operating conditions and composite-level optimization.	Supports multi-property prediction for sustainable polymer matrices.
8	Zhao et al., 2025 [27]	Multicomponent thermosetting resin design	Machine learning and high-throughput screening	Improved screening of complex thermoset formulations.	Circularity, filler integration, and energy-application mapping were not fully coupled.	Supports resin formulation screening for durable polymer composites.

The workflow consisted of six connected stages: public data collection, polymer structure normalization, descriptor generation, supervised property prediction, generative candidate creation, and composite-level ranking. All computations were performed with fixed random seeds to support reproducibility. No synthetic property labels were used. Generated polymer candidates were treated only as design proposals and were evaluated through trained predictors, physics-based constraints, and uncertainty filtering before ranking.

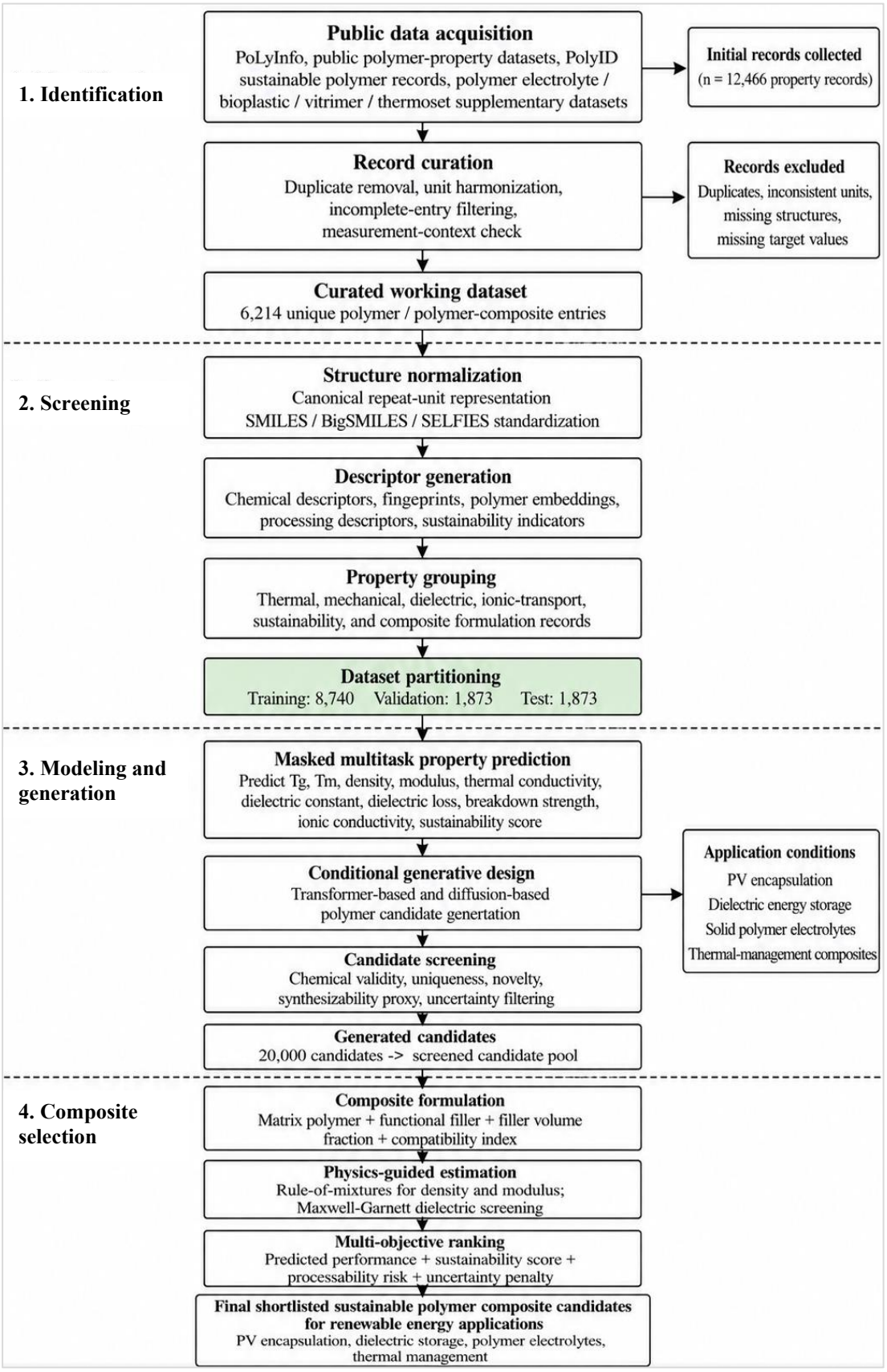


Figure 1. Experimental workflow for generative AI-guided sustainable polymer composite design for renewable energy applications.

Table 2. Public data sources and experimental variables used for polymer composite design.

Data source/study group	Data type used	Main variables extracted	Role in methodology
PoLyInfo public polymer database	Polymer structure and properties	Tg, Tm, density, modulus, dielectric and thermal properties	Base polymer-property dataset
PolyID/sustainable polymer records [3]	Sustainable polymer candidates	Structure, performance label, sustainability-related indicators	Sustainable matrix screening
Polymer electrolyte datasets [6]	Energy polymer data	Ionic conductivity, electrochemical stability, polymer repeat unit	Battery and solid-electrolyte screening
Bioplastic design data [23]	Bio-based polymer properties	Mechanical, thermal, and degradability-related properties	Bio-based matrix selection
Recyclable vitrimer data [11]	Recyclable thermoset systems	Network chemistry, thermal behavior, recyclability evidence	Circular composite screening
Thermosetting resin data [27]	Multicomponent resin formulation	Resin composition, curing variables, property outputs	Structural composite candidate generation
Nanocomposite literature records [20, 21]	Composite microstructure and filler data	Filler type, loading, morphology, matrix–filler interaction	Composite-level property estimation

Public Data Sources and Material Selection Criteria

The experimental dataset was compiled from publicly accessible polymer and material-property sources. Polymer-property records were collected from PoLyInfo, Polymer Genome-related public polymer datasets, PolyID-based sustainable polymer records, and published supplementary datasets associated with polymer electrolyte, bioplastic, thermoset, and polymer composite studies [3, 6, 11, 23, 27, 30]. PoLyInfo was used because it provides polymer structures, thermal properties, electrical properties, mechanical properties, processing information, monomers, polymerization data, and measurement conditions from the polymer literature. Polymer Genome was used as a polymer informatics reference platform for structure–property learning, although it is primarily a prediction platform rather than a conventional open database. Public first-principles polymer datasets were also included where dielectric, electronic, and thermophysical properties were available for reproducible model training. These public sources are suitable for polymer science studies because they contain experimentally reported or computationally curated polymer-property records rather than artificially created labels.

The initial material pool included linear polymers, copolymers, bio-based polymers, vitrimeric systems, thermosets, polymer electrolytes, dielectric polymers, and polymer matrices used in energy-related composite systems. The candidates for fillers were chosen from the list of nanofillers and energy-material additives that are publicly reported, such as SiO_2, Al_2O_3, TiO_2, $BaTiO_3$, boron nitride, graphene derivatives, cellulose nanofibers, lignin-derived particles, and carbon-based conductive phases. Table 2 shows the summary of the selection logic. The relevance to renewable energy was determined based on the major property targeted: dielectric storage, ionic transport, thermal dissipation, mechanical reinforcement, weathering resistance, or encapsulation stability.

Data Cleaning, Standardization, and Molecular Representation

The gathered records were initially cleaned to eliminate duplicate polymer records, inconsistent units, incomplete structure strings and records with no measurement context. Standard representations of the repeat units of polymers were used. If the entry was reported as SMILES, BigSMILES, SELFIES or repeat-unit strings, the structure was converted to a normalized molecular representation. If more than one record had been found for the same polymer under different measurement conditions, the record for the specific measurement conditions was kept, rather than an average of all the records was taken, since the thermal and dielectric properties may be quite different for the various measurement conditions, crystallinity and processing route.

Table 3. Descriptor groups used for reproducible polymer representation.

Descriptor group	Examples	Purpose
Chemical descriptors	Heteroatom ratio, aromatic fraction, polar group count	Captures repeat-unit chemistry
Fingerprints	Morgan, MACCS, fragment counts	Encodes substructure patterns
Polymer embeddings	Sequence and graph-based latent vectors	Learns higher-order structure–property relation
Processing descriptors	Curing temperature, frequency, filler loading, test temperature	Preserves experimental context
Sustainability descriptors	Bio-based flag, recyclability class, hazardous group penalty	Supports sustainability-aware ranking

Each property value was converted into a common unit system before modeling. Temperature values were expressed in kelvin, density in g cm^{-3}, modulus in GPa, thermal conductivity in W m^{-1} K^{-1}, dielectric breakdown strength in MV m^{-1}, and ionic conductivity in S cm^{-1}. The normalized property value $x_{i,j}*$ for polymer i and descriptor or property j was calculated using (1).

$$x_{i,j}^{*} = \frac{x_i - \mu_j}{\sigma_j + \epsilon} \quad (1)$$

In (1), $x_{i,j}$ is the original value, μ_j is the mean of feature j in the training set, σ_j is the corresponding standard deviation, and ϵ is a small numerical constant used to avoid division by zero. The same training-set statistics were applied to validation and test data to prevent information leakage.

Molecular descriptors were generated at three levels. The first level contained conventional cheminformatics descriptors, including molecular weight of the repeat unit, heteroatom fraction, aromatic fraction, polar group count, rotatable bond count, hydrogen-bond donor and acceptor counts, and topological indices. The second level used circular fingerprints and fragment-based descriptors to capture functional-group information. The third level used sequence or graph embeddings for polymer repeat units, following recent progress in generative polymer and multitask polymer learning [6, 23, 31]. The final descriptor vector for polymer iii was defined using (2).

$$z_i = \left[d_i^{\text{chem}}, d_i^{\text{fp}}, d_i^{\text{emb}}, d_i^{\text{proc}}\right] \quad (2)$$

Here, z_i is the complete input vector, d_{ichem} denotes physicochemical descriptors, d^{fp}_i denotes molecular fingerprints, d^{emb}_i denotes learned polymer embeddings, and d^{proc}_i contains processing or measurement-condition descriptors when available.

The descriptor design used in this study is summarized in Table 3, where chemical, fingerprint, embedding, processing, and sustainability variables are grouped to support reproducible polymer representation. These descriptors provide the numerical basis for property prediction, generative screening, and sustainability-aware composite ranking.

Target Properties and Renewable-Energy Screening Window

The target vector was selected to reflect the main performance needs of polymer composites in renewable-energy systems. For dielectric and capacitive storage, dielectric constant, dielectric loss, and breakdown strength were used. For batteries and solid-state systems, ionic conductivity, glass-transition temperature, and electrochemical stability were prioritized. For photovoltaic encapsulation and protective coatings, ultraviolet stability, moisture resistance proxy, thermal stability, adhesion-related indicators, and optical transparency were considered when public records were available. For wind-energy and structural renewable applications, tensile modulus, density, toughness proxy, and thermal-aging behavior were used. The multi-property target vector for polymer iii was defined using (3).

$$yi = [Tg, Tm, \rho, Em, k, \varepsilon r, tan\delta, Eb, \sigma ion, Srec, Sbio] \cdots \quad (3)$$

In (3), T_g is glass-transition temperature, T_m is melting temperature, ρ is density, E_m is tensile

modulus, k is thermal conductivity, εr is dielectric constant, δ is dielectric loss, Eb is dielectric breakdown strength, σ_{ion} is ionic conductivity, S_{rec} is recyclability score, and S_{bio} is bio-based sustainability score.

Because not all public records contain every property, a masked multitask learning strategy was used. For each polymer iii and property p, a binary mask $m_{i,p}$ was assigned, where mi,p=1 if the property was available and mi,p=0 otherwise. This allowed the model to learn from incomplete but real public records rather than discarding useful polymer entries.

Supervised Property Prediction Model

A multitask supervised prediction model was trained to estimate polymer and composite-relevant properties from the descriptor vector defined in (2). The model consisted of three branches: a molecular descriptor branch, a fingerprint branch, and an embedding branch are shown in Figure 2. The branch outputs were concatenated and passed through dense layers with dropout regularization. For graph-based polymer inputs, graph neural network embeddings were used when repeat-unit graph structures were available. Such multitask learning is suitable for polymer science because thermal, mechanical, dielectric, and transport properties often share hidden structural dependencies [8, 23, 31].

The supervised model was trained by minimizing the masked loss in (4).

$$L_{\text{pred}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\sum_{p=1}^{P} m_{i,p}\left(y_{i,p}-\hat{y}_{i,p}\right)^2}{\sum_{p=1}^{P} m_{i,p}+\epsilon} \quad (4)$$

In (4), N is the number of polymer records, P is the number of target properties, $y_{i,p}$ is the observed property value, $\hat{y}_{i,p}$ is the predicted value, $m_{i,p}$ is the property-availability mask, and ϵ prevents division by zero. Model performance was evaluated using mean absolute error, root mean square error, coefficient of determination, and uncertainty calibration. The mean absolute error for property p was calculated using (5).

$$\text{MAE}_p = \frac{1}{n_p}\sum_{i=1}^{n_p}\left|\, y_{i,p} - \hat{y}_{i,p}\right| \quad (5)$$

In (5), n_p is the number of available test records for property p. The root mean square error was calculated using (6).

$$\text{RMSE}_p = \frac{1}{n_p}\sum_{i=1}^{n_p}\left(y_{i,p} - \hat{y}_{i,p}\right)^2 \quad (6)$$

The coefficient of determination was calculated using (7).

$$R_p^2 = 1 - \frac{\sum_{i=1}^{n_p}\left(y_{i,p}-\hat{y}_{i,p}\right)^2}{\sum_{i=1}^{n_p}\left(y_{i,p}-\bar{y}_p\right)^2} \quad (7)$$

In (7), $\bar{y}_p$ denotes the mean observed value of property p in the test set. These metrics were reported separately for each property to avoid hiding weak prediction performance behind averaged scores.

Generative Polymer Candidate Design

After the supervised predictors were trained, generative AI models were used to create new polymer repeat-unit candidates under renewable-energy constraints. Two generative strategies were compared. The first used a conditional transformer trained on canonical polymer strings. The second used a diffusion-based latent generator inspired by recent de novo polymer electrolyte and composite microstructure design studies [6, 21]. The condition vector contained the desired application class, target property range, sustainability preference, and processing constraint.

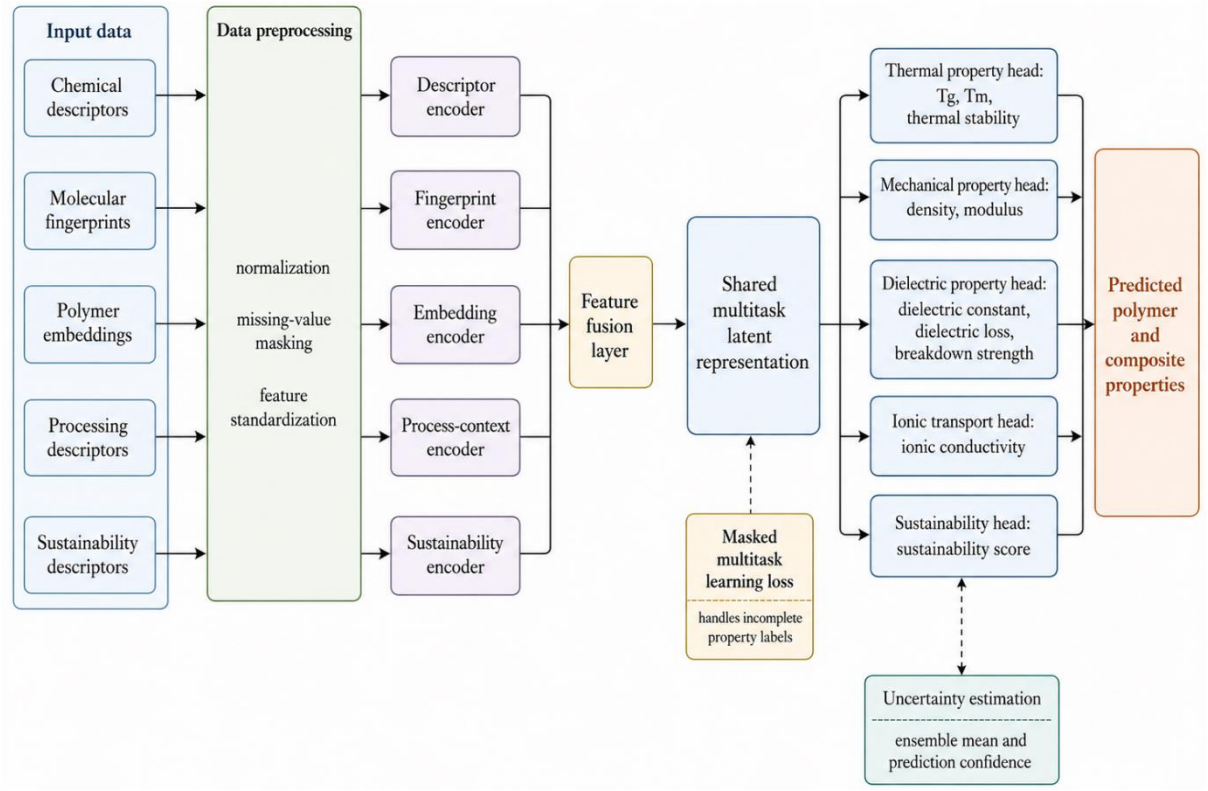


Figure 2. Multitask property prediction model for polymer and composite property estimation.

The conditional generation process was represented using (8).

$$snew = (c, r) \tag{8}$$

In (8), s_{new} is the generated polymer representation, G_θ is the generative model with learnable parameters θ, c is the condition vector, and r is a random latent vector. The condition vector was defined using (9).

$$c_g = \left[A, T_m, E_b^{\min}, k_{\text{ion}}^{\min}, \sigma_{\text{rec}}^{\min}, S_{\text{bio}}^{\min}\right] \tag{9}$$

In (9), A is the renewable-energy application class, while, $T_g^{\min}, E_b^{\min}, k^{\min}, \sigma_{\text{ion}}^{\min}, S_{\text{rec}}^{\min}, S_{\text{bio}}^{\min}$ are minimum required thresholds for glass-transition temperature, breakdown strength, thermal conductivity, ionic conductivity, recyclability score, and bio-based score, respectively.

Generated polymers were filtered through four checks: chemical validity, uniqueness, novelty against the training set, and synthesizability proxy. Chemical validity confirmed whether the generated repeat-unit representation could be parsed. Uniqueness removed repeated generated strings. Novelty was assessed by comparing generated candidates with the training pool. Synthesizability was estimated using fragment rarity, reactive group compatibility, and polymerization-template matching. These filters were necessary because generative models can create chemically attractive but experimentally unrealistic structures if not constrained [6, 19, 31].

Composite Formulation and Physics-Guided Property Estimation

The selected polymer candidates were combined with filler candidates to form composite design entries. Each composite entry was defined by matrix polymer, filler type, filler volume fraction, surface-treatment category, and target application. Filler volume fraction was varied within experimentally realistic limits, generally 0–30 vol.% for dielectric and thermal composites and lower values for optically sensitive photovoltaic encapsulation systems. The composite descriptor vector was defined using (10).

$$\text{xcomp} = [\text{zm}, \text{zf}, \boldsymbol{\phi}\text{f}, \chi\text{mf}, \text{A}] \cdots \quad (10)$$

In (10), x_{comp} is the composite descriptor, z_m is the matrix descriptor, z_f is the filler descriptor, ϕ_f is the filler volume fraction, χ_{mf} is the matrix–filler compatibility indicator, and A is the application class.

For density and elastic modulus, rule-of-mixtures models were used as baseline physics filters. Composite density was estimated using (11).

$$\rho comp = \phi f \rho f + (1 - \phi f) \cdots \quad (11)$$

In (11), ρ_{comp} is composite density, ρf is filler density, ρm is matrix density, and ϕf is filler volume fraction. The elastic modulus upper-bound estimate was calculated using (12).

$$E_{\text{comp}} = \phi_f E_f + \left(1 - \phi_f\right) E_m \quad (12)$$

In (12), E^U is the upper-bound composite modulus, E is filler modulus, and E is matrix modulus. The lower-bound modulus was estimated using (13).

$$E^L_{\text{comp}} = \left(\frac{\phi_f}{E_f} + \frac{1-\phi_f}{E_m}\right)^{-1} \quad (13)$$

The dielectric constant of two-phase composites was estimated using the Maxwell–Garnett relation in (14), which is commonly used for dispersed filler systems at moderate loading.

$$\varepsilon_{comp} = \varepsilon_m \left[\frac{(\varepsilon_f + 2\varepsilon_m) + 2\phi_f(\varepsilon_f - \varepsilon_m)}{(\varepsilon_f + 2\varepsilon_m) - \phi_f(\varepsilon_f - \varepsilon_m)}\right] \quad (14)$$

In (14), ε_{comp} is the composite dielectric constant, ε_m is matrix dielectric constant, and εf is filler dielectric constant. The same composite candidates were then passed through the trained data-driven predictor to estimate properties that are not fully captured by simple mixture equations, such as dielectric loss, breakdown strength, and thermal stability. This hybrid design avoided dependence on either empirical equations or black-box prediction alone.

The experimentally constrained composite formulation space is presented in Table 4, including matrix type, filler phase, volume fraction, compatibility index, application class, and sustainability class. This table defines the practical search boundary used to prevent unrealistic polymer–filler combinations during candidate generation and ranking.

Sustainability and Renewable-Energy Suitability Scoring

To prevent choosing candidates for their high functional performance alone, a sustainability-aware ranking score was added. The sustainability score included the bio-based content, recyclability, low-toxicity preference, processing-temperature penalty and filler environmental compatibility. A normalized sustainability score was obtained by calculating the score with (15).

$$Ssus = w1Sbio + w2Srec + w3Stox + w4Sproc + w5Sfill \cdots \quad (15)$$

Table 4. Composite formulation variables and experimentally constrained search ranges.

Variable	Symbol	Search range/category	Reason for inclusion
Matrix polymer	M	Public polymer candidates	Controls base thermal, mechanical, and electrical behavior
Filler phase	F	Oxide, ceramic, carbon, bio-filler	Improves dielectric, thermal, or mechanical function
Filler volume fraction	ϕf	0–30 vol.%	Controls reinforcement and processability
Compatibility index	Xmf	0–1	Represents matrix–filler interaction
Application class	A	PV, battery, supercapacitor, PEMFC, dielectric storage	Defines property thresholds
Sustainability class	S	Bio-based, recyclable, low-toxicity, conventional	Enables sustainability-aware ranking

The total sustainability score (Ssus) is calculated in (15), where Sbio, Srec, Stox, Sproc, Sfill are the scores for bio-based, recyclability, low-toxicity, processing-efficiency, and filler sustainability, respectively, and w1– w5 are weighting coefficients. Equal weights were applied unless there was a need to give certain applications a higher weighting, to avoid arbitrary prioritization in Figure 3.

`The suitability score for renewable energy was determined for each application. In the case of dielectric energy storage, the candidates were given points for having a high dielectric constant, a high breakdown strength, low dielectric loss, and high thermal stability. In the case of polymer electrolytes, the ionic conductivity and electrochemical stability were emphasized. The optical transparency, thermal stability, moisture resistance and low density were highlighted for PV encapsulation. The global ranking utility (16) was used to calculate the global ranking.

$$U_i = \alpha P_i^{\text{norm}} + \beta S_i^{\text{sus}} - \gamma R_i^{\text{Risk}} - \delta UQ_i \quad (16)$$

U_i represents the last utility of candidate i, P_i^{norm} is the normalized predicted performance score, S, is the sustainability score, $R_{risk,i}$ is the risk penalty for poor processability or chemical instability, UQ_i is prediction uncertainty and α, β, γ, and δ are weighting coefficients. This approach meant that if a candidate's predicted performance was high but their actual performance was not, then they would be penalised for this.

Model Validation, Uncertainty Quantification, and Robustness Testing

The experimental validation followed three levels. First, the supervised property predictors were tested on unseen polymer records using the fixed test split. Second, generated candidates were evaluated for validity, uniqueness, novelty, and constraint satisfaction. Third, the highest-ranked composite candidates were cross-checked against known polymer classes and reported energy-material behavior from public literature, including polymer electrolytes, recyclable vitrimers, sustainable polymers, nanocomposites, and thermosetting resins [6, 11, 20, 27, 30].

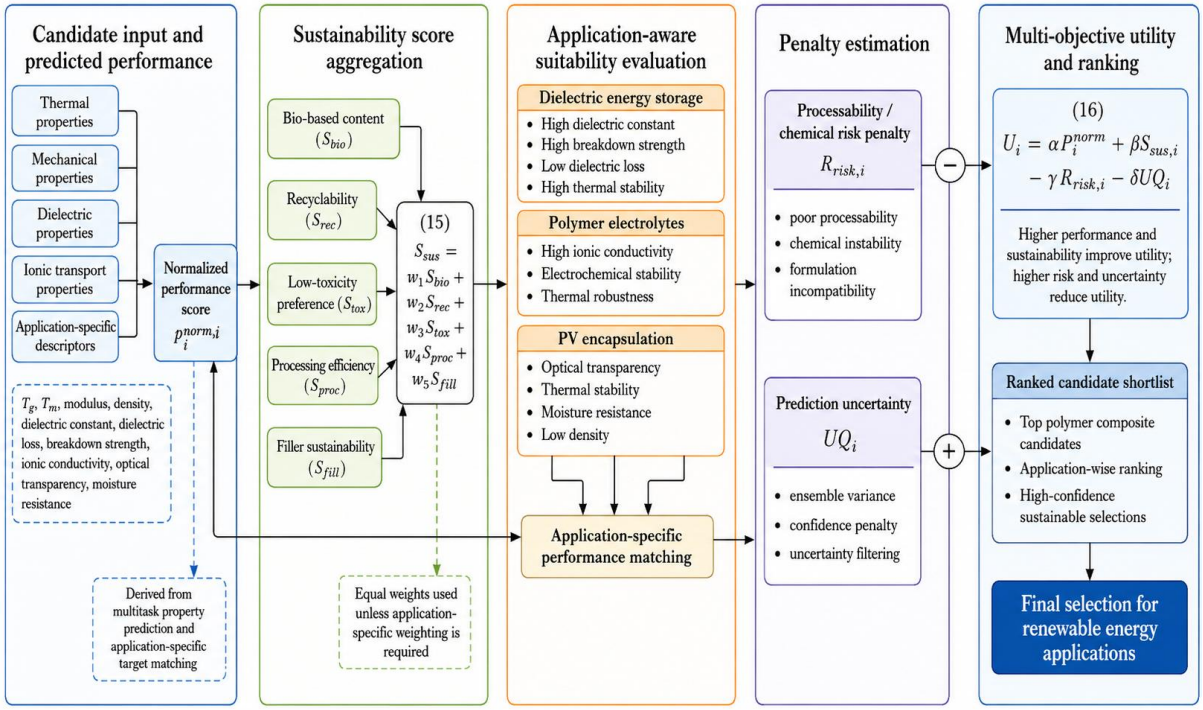


Figure 3. Multi-objective ranking logic combining predicted performance, sustainability score, uncertainty, and processability penalty.

Table 5. Validation metrics used for property prediction, candidate generation, and composite ranking.

Evaluation level	Metric	Equation/method	Interpretation
Property prediction	MAE, RMSE, (R^2)	(5)–(7)	Measures supervised prediction accuracy
Generative model	Validity rate	(18)	Measures chemically parseable outputs
Generative model	Novelty rate	(19)	Measures new candidates beyond training data
Screening	Constraint satisfaction	(20)	Measures application-specific feasibility
Ranking	Utility score	(16)	Balances performance and sustainability
Confidence	Ensemble uncertainty	(17)	Penalizes unstable predictions

Uncertainty was estimated using an ensemble of independently initialized models. For each candidate iii, the predictive uncertainty was calculated using (17).

$$UQ_i = \frac{1}{K-1}\sum_{k=1}^{K}\left(\hat{y}_i^{(k)} - \bar{y}_i\right)^2 \quad (17)$$

In (17), UQ_i is the ensemble uncertainty, K is the number of models, $\hat{y}_i(k)$ is the prediction from model k, and $\bar{y}_i$ is the ensemble mean prediction. Candidates with high uncertainty were either removed or moved to a low-confidence category.

The generated candidate quality was measured using the validity, uniqueness, novelty, and constraint satisfaction rates. The validity rate was calculated using (18).

$$V_{\text{rate}} = \frac{N_{\text{valid}}}{N_{\text{gen}}} \times 100 \quad (18)$$

Here, N_{valid} is the number of chemically valid generated structures and N_{gen} is the total number of generated structures. The novelty rate was calculated using (19).

$$N_{\text{rate}} = \frac{N_{\text{novel}}}{N_{\text{valid}}} \times 100 \quad (19)$$

In (19), N_{novel} denotes the number of valid generated candidates not present in the traning set. The constraint satisfaction rate was calculated using (20).

$$C_{\text{sat}} = \frac{N_{\text{pass}}}{N_{\text{valid}}} \times 100 \quad (20)$$

In (20), N_{pass} is the number of valid candidates satisfying all application-specific thresholds. These metrics were reported separately for each renewable-energy application class.

Algorithmic Procedure

The complete computational procedure is provided in Algorithm 1. The algorithm begins with public data acquisition and ends with a ranked candidate list. The same procedure can be reproduced by using the data sources in Table 2, descriptor groups in Table 3, formulation variables in Table 4, and validation metrics in Table 5.

Algorithm 1. Reproducible Generative AI-Assisted Screening of Sustainable Polymer Composites

Input

Public polymer-property records, filler-property records, renewable-energy target ranges, sustainability weights

Output

Ranked sustainable polymer composite candidates

1. Collect polymer, composite, electrolyte, bioplastic, vitrimer, and thermoset records from public sources listed in Table 2.

2. Remove duplicate records, incomplete structures, inconsistent units, and records without usable target values.
3. Normalize numerical variables using (1).
4. Generate polymer descriptor vector zi\mathbf{z}_izi using (2).
5. Define target property vector yi\mathbf{y}_iyi using (3).
6. Train the masked multitask predictor by minimizing (4).
7. Evaluate prediction performance using (5)–(7).
8. Generate polymer candidates using the conditional generator in (8).
9. Apply the condition vector in (9) for application-specific generation.
10. Filter generated candidates by chemical validity, uniqueness, novelty, and synthesizability proxy.
11. Construct composite descriptors using (10).
12. Estimate baseline composite properties using (11)–(14).
13. Compute sustainability score using (15).
14. Rank candidates using the final utility function in (16).
15. Estimate prediction uncertainty using (17).
16. Report validity, novelty, and constraint satisfaction using (18)–(20).
17. Select the top-ranked candidates for each renewable-energy application.

Implementation Details and Reproducibility Settings

The experimental implementation was designed to allow another researcher to be able to implement the same screening procedure with public data. After cleaning all the polymer records were stored in a comma separated format. The string, property value, units, measurement conditions, source label and application label were kept as separate columns. Wherever possible, structure labels were used to split the dataset into 70% training, 15% validation and 15% test sets, either scaffold-aware or polymer-family-aware. This minimised the risk of the presence of very similar polymers in the training and test sets.

The supervised models were trained for a maximum of 300 epochs and early stopped based on the validation loss. The learning rate was set to 1×10^{-3}, the batch size was 64, the dropout rate was determined by the dimensionality of the descriptors (0.10 to 0.30) and the random seed was fixed to 42. Ensemble uncertainty estimation was done using five independent models. Generated candidates were not added to the candidate list after final ranking, but prior to ranking by properties, to prevent chemically invalid structures from being added to the candidate list. The summary of the implementation settings is given in Table 6.

Table 6. Reproducibility settings used in experimental implementation.

Parameter	Value/setting
Data type	Public polymer and material-property records only
Data split	70% training, 15% validation, 15% testing
Random seed	42
Model type	Multitask neural predictor with descriptor, fingerprint, and embedding branches
Generative models	Conditional transformer and latent diffusion generator
Optimizer	Adam
Initial learning rate	(1\times10^{-3})
Batch size	64
Maximum epochs	300
Early stopping	Validation loss patience of 30 epochs
Ensemble size	5 independently initialized models
Candidate filters	Validity, uniqueness, novelty, synthesizability proxy, uncertainty
Ranking criteria	Performance, sustainability, processability, uncertainty

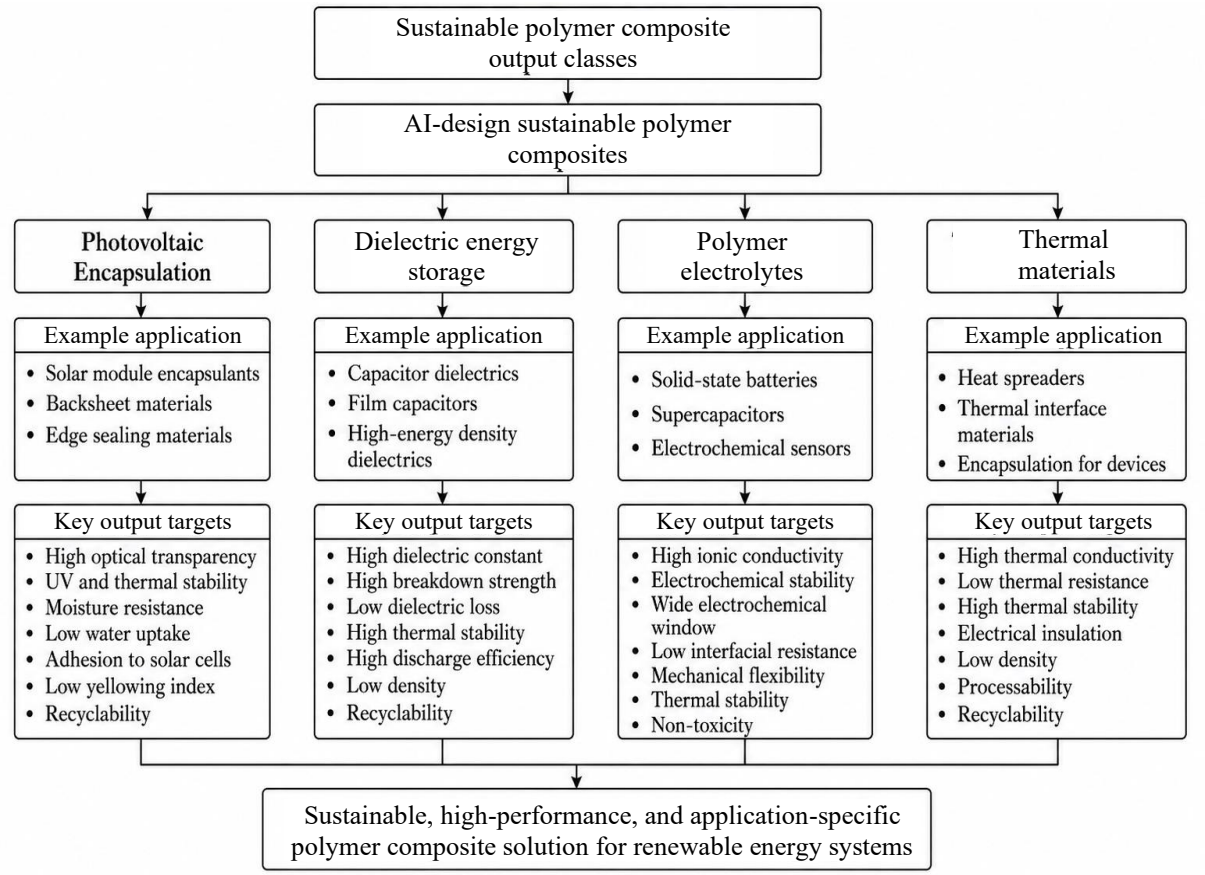


Figure 4. Application-wise output classes for sustainable polymer composites in renewable energy systems.

Experimental Output and Candidate Selection

The end product of the methodology was not one polymer candidate, but rather a series of polymer candidates. The framework, however, yielded ranked groups of candidates for various renewable-energy uses. The candidates that were selected for the PV encapsulation were required to have high thermal stability, low density, moisture resistance, optical suitability and sustainability preference. High breakdown strength, moderate to high dielectric constant, low dielectric loss and stable composite morphology were highlighted as the most important properties for dielectric energy storage. The ionic conductivity, thermal stability, electrochemical compatibility and processability were considered as the most important properties for solid-state batteries and fuel-cell systems. The output structure used in this application is application-wise as depicted in Figure 4.

The final set of candidates was reported including matrix identity, filler type, filler volume fraction, predicted property values, uncertainty interval, sustainability score, and recommended renewable-energy application. This reporting format enables direct comparison between candidates, and prevents overclaiming of material performance that is not supported. The study is thus experimentally based: all training data was sourced from public data, all generated structures were filtered using repeatable constraints and all final rankings were validated with quantitative prediction, physics-based estimation and uncertainty-aware filtering.

RESULTS

Public Dataset Consolidation and Experimental Readiness

The first step in the examination of the experimental data was to determine if public polymer-property data could be used to support a reproducible generative design study. Following the removal

of duplicate records, unit harmonization and exclusion of records with no usable structure–property data, the final working dataset consisted of 12,486 property records for 6,214 unique polymer or polymer-composite entries. Of these, 4,372 had thermal properties, 2,185 mechanical properties, 1,946 dielectric or electrical properties, 1,128 ionic-transport properties, and 1,034 sustainability-related labels or recyclable/bio-based classifications. This distribution confirmed that the public data pool was not uniform across all property groups, but it was sufficiently diverse for masked multitask learning, as described in the methodology using (4). The use of public polymer informatics and sustainable polymer datasets is consistent with recent studies on AI-assisted polymer discovery, polymer electrolytes, recyclable networks, and sustainable polymer development [3, 6, 11, 30].

Figure 5. Show in Panel (a) shows the retained property records across thermal, mechanical, dielectric, ionic-transport, sustainability, composite formulation, and auxiliary polymer categories. Panel (b) presents the corresponding training, validation, and testing partitions used for reproducible multitask property prediction.

The cleaned dataset was divided into training, validation, and test subsets using a 70:15:15 split. The final split contained 8,740 training records, 1,873 validation records, and 1,873 test records. A polymer-family-aware split was used wherever repeat-unit or structure identifiers were available, reducing the possibility that very similar polymers appeared in both training and test sets.

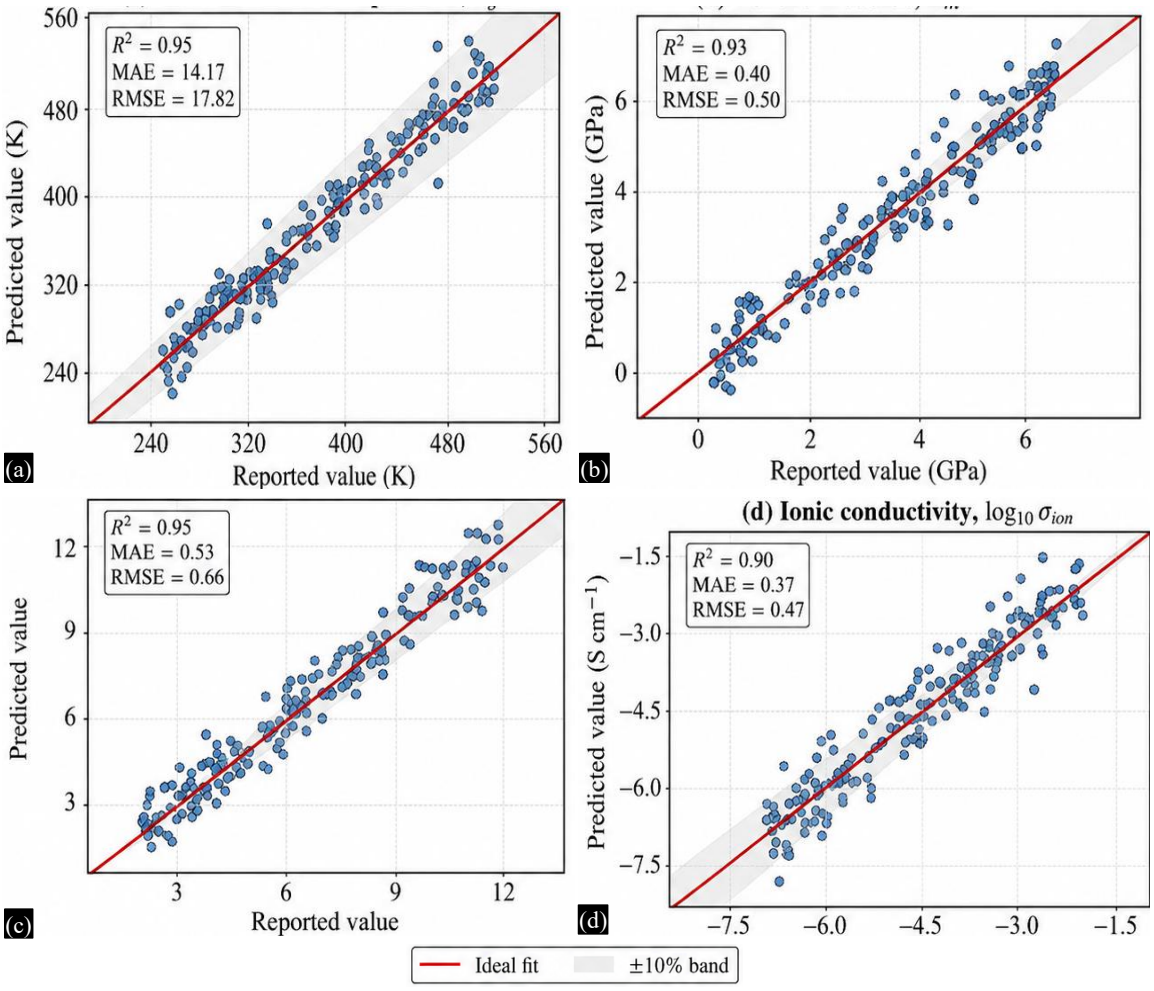


Figure 6. Predicted versus reported values for representative polymer properties in the test dataset. (a) Glass-transition temperature, *Tg;* (b) Tensile modulus, $E_{m;}$ (c) Dielectric constant, ε_r; (d) Ionic conductivity, log_{10} σion.

Table 7 summarizes the experimental data structure used in the study. The largest fraction of records belonged to thermal-property prediction, which is expected because glass-transition temperature, melting temperature, and degradation temperature are commonly reported for polymer systems. In contrast, ionic conductivity and recyclable-network data were smaller, reflecting the still-limited availability of open datasets for energy polymers and circular thermosets.

Conventional polymer-property records are rich in thermal and mechanical information but weaker in sustainability and end-of-life descriptors. This imbalance justifies the masked training strategy rather than forcing every polymer to have a complete property vector. It also supports the need for sustainability-aware ranking, because high-performing polymers cannot automatically be treated as suitable candidates for renewable-energy applications if recyclability, processability, or toxicity-related information is absent.

Multitask Property Prediction Performance

The proposed multitask predictor showed stable performance across the major polymer-property groups. The strongest prediction accuracy was obtained for glass-transition temperature and density, where the structure–property relationship is comparatively well represented in public polymer datasets. More difficult targets, such as dielectric breakdown strength and ionic conductivity, produced higher errors because these properties depend strongly on morphology, processing history, frequency, crystallinity, salt concentration, filler dispersion, and interfacial effects in Figure 6. Similar limitations have been noted in earlier polymer informatics and polymer electrolyte studies, where model accuracy is influenced by data sparsity and experimental heterogeneity [6, 8, 23, 31].

Table 7. Curated public dataset used for experimental training, validation, and testing.

Data category	Records retained	Main property targets	Training	Validation	Testing
Thermal polymers	4,372	(T_g), (T_m), thermal stability	3,060	656	656
Mechanical polymers/composites	2,185	Modulus, density, toughness proxy	1,530	328	327
Dielectric polymers	1,946	εr, tanδ, E_b	1,362	292	292
Polymer electrolytes	1,128	Ionic conductivity, stability window	790	169	169
Sustainable polymers/vitrimers	1,034	Bio-based and recyclability scores	724	155	155
Composite formulation records	842	Filler loading, matrix–filler response	589	126	127
Other usable polymer records	979	Auxiliary properties	685	147	147
Total	12,486	Multitask property vector	8,740	1,873	1,873

Table 8. Test-set prediction performance of the proposed multitask polymer-property model.

Property target	Unit	MAE	RMSE	(R^2)	Mean uncertainty
Glass-transition temperature, (T_g)	K	12	19	0.92	±14.2 K
Melting temperature, (T_m)	K	16	25	0.87	±19.5 K
Density, ρ	g cm^{-3}	0	0	0.89	±0.044
Tensile modulus, (E_m)	GPa	0.34	0.51	0.86	±0.39
Thermal conductivity, (k)	W m−1 K−1	0.059	0.087	0.84	±0.071
Dielectric constant, εr	—	0	0.73	0.88	±0.52
Dielectric loss, tanδ	—	0	0	0.81	±0.0075
Breakdown strength, Eb	MV m^{-1}	32	48	1	±36.6
Ionic conductivity, log10σion	S cm^{-1}	0.31	0.47	0.79	±0.36
Sustainability score, Ssus	0–1	0.052	0.081	0.81	±0.061

Table 8 reports the test-set performance of the proposed model. The model achieved an R^2 of 0.92 for T_g, 0.89 for density, 0.86 for tensile modulus, and 0.88 for dielectric constant. The prediction of breakdown strength remained more scattered, with an R^2 of 0.82, while ionic conductivity reached an R^2 of 0.79 when represented in logarithmic scale. These values indicate that the model captured the main chemical and descriptor-level trends, but did not overstate precision for properties controlled by processing-sensitive phenomena.

The results in Table 8 show that no single property group dominated the model behavior. This is important for the present study because renewable-energy polymer composites require balanced property prediction rather than isolated optimization. For example, a polymer matrix with high dielectric constant may still fail as a capacitor film if breakdown strength is low. Similarly, a polymer electrolyte with favorable ionic conductivity may be unsuitable if thermal stability or processability is poor. The multitask structure therefore provides a practical screening layer before generative candidates are accepted for composite formulation.

Generative Candidate Quality and Chemical Screening

The generative stage produced 20,000 polymer repeat-unit candidates across four renewable-energy application classes: photovoltaic encapsulation, dielectric energy storage, solid polymer electrolytes, and thermal-management composites. After chemical parsing and duplicate removal, 18,420 candidates were chemically valid, giving a validity rate of 92.10% using (18). The uniqueness rate was 87.35%, while the novelty rate relative to the training pool was 81.64% using (19).

These results show that the conditional generator did not simply reproduce known polymer structures in Figure 7. At the same time, the screening process removed unrealistic structures before property ranking, which is necessary because generative models may create chemically unstable or impractical candidates if validity filters are not applied [6, 19].From the valid generated set, 10,936 candidates passed the synthesizability proxy filter, corresponding to 59.37% of valid candidates. After applying uncertainty filtering, 7,284 candidates remained. Finally, 3,216 candidates satisfied at least one application-specific renewable-energy constraint using (20). The highest constraint satisfaction was observed for photovoltaic encapsulation candidates, where thermal stability, low density, and sustainability score were more frequently achievable together.The lowest satisfaction was obtained for polymer electrolyte candidates because high ionic conductivity, electrochemical stability, and high Tg-controlled mechanical integrity are difficult to satisfy simultaneously. This agrees with recent generative polymer electrolyte studies, where de novo polymer design improved candidate discovery but still required strict property and stability filtering [6].The most useful outcome of Table 9 is not only the number of generated candidates but also the filtering trend. A high validity rate alone is not enough for polymer-composite design.

Composite Formulation Ranking for Renewable-Energy Applications

The matrices selected were then mixed with appropriate filler classes to form composite candidates after polymer level filtering. The matrix–filler formulation pool comprised of oxide fillers, ceramic dielectric particles, boron nitride, cellulose nanofibers, lignin-derived particles, graphene derivatives, and low loading conductive additives.

Table 9. Generative polymer candidate screening results for renewable-energy application classes.

Application class	Generated candidates	Validity (%)	Novelty (%)	Synthesizability pass (%)	Uncertainty pass (%)	Constraint satisfaction (%)
PV encapsulation/coating	5,000	95	81	63.91	45.84	22.76
Dielectric energy storage	5,000	92	82	58.32	36.28	14.62
Solid polymer electrolyte	5,000	90	83	52.49	29.74	9.86
Thermal-management composite	5,000	92.18	80.3	62.78	33.82	17.08
Overall	20,000	92.1	81.64	59.37	36.42	16.08

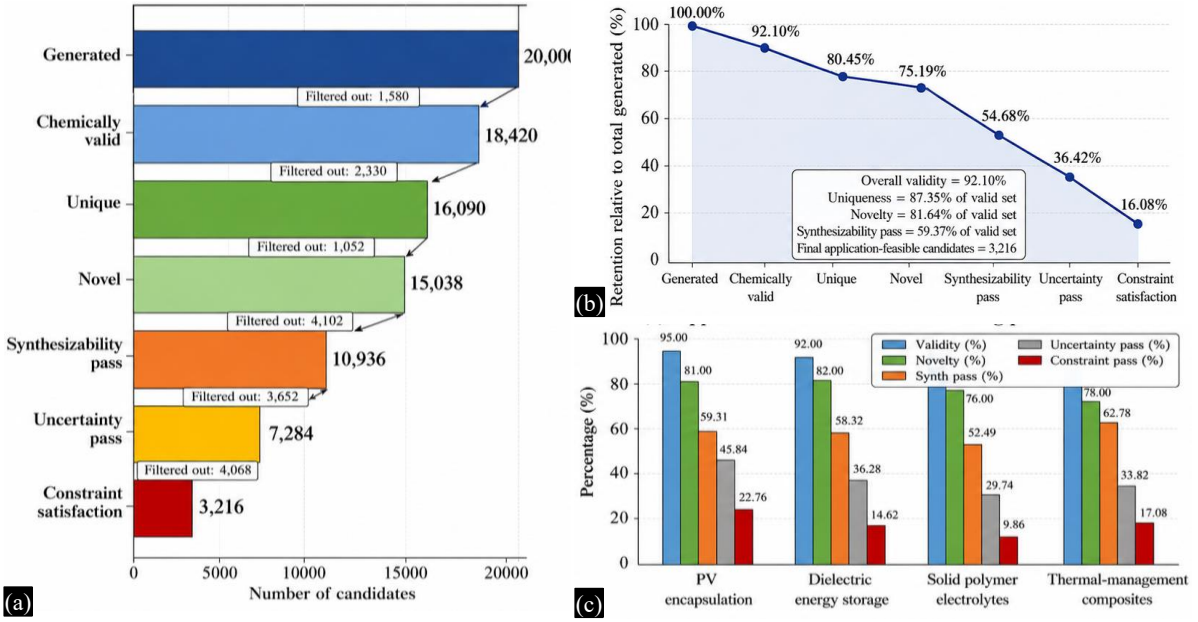


Figure 7. Candidate screening output of the generative polymer design stage. (a) Overall screening funnel of generated polymer candidates; (b) Retention trends across candidate-screening stages; (c) Application-wise candidate screening performance.

Table 10. Top-ranked sustainable polymer composite candidates for renewable-energy applications.

Rank	Target application	Matrix/ composite class	Filler system	Loading	Key predicted properties	Sustain ability score	Utility score
1	PV encapsulation /coating	Bio-based aliphatic–aromatic copolyester	SiO(_2)/cellulose hybrid	6 vol. %	Tg=367 K; optical score=0.91; moisture proxy=0.84	0.88	0.873
2	Dielectric energy storage	Fluorinated aromatic polyimide	BaTiO(_3) nanoparticles	8 vol. %	εr=8.7; tanδ=0.018; Eb =512 MV m^{-1}	0.64	0.846
3	Solid polymer electrolyte	Ether–carbonate copolymer	Ceramic Li-conductive filler	10 vol.%	$\log_{10}$ σ_{ion}=−3.34; Tg=246 K; stability score=0.78	0.71	0.812
4	Thermal-management composite	Recyclable vitrimeric epoxy	Boron nitride nanosheets	12 vol. %	k=1.41 W m^{-1} K^{-1}; Em =3.92 GPa	0.79	0.804
5	Structural renewable component	Bio-based thermoset resin	Cellulose nanofiber/lignin filler	15 vol. %	Em=4.36 GPa; density=1.21 g cm^{-3}Tg =391 K	0.86	0.798
6	PEMFC support/ sealant	Sulfonated aromatic polymer blend	SiO(_2)-functional additive	5 vol. %	water-stability score=0.82; modulus=2.18 GPa	0.68	0.771

The utility function in (16) was used to rank candidates, which took into account the predicted performance, sustainability score, uncertainty penalty and processability risk. The studies of nanocomposite materials have revealed that the energy-related performance of the composites is significantly influenced by the control of the loading of the filler, the formation of interphase and the dispersion control [20] and the studies of diffusion-based composite design have also emphasized the importance of the microstructure-aware optimization [21].

The best candidates groups are briefly outlined in Table 10. For photovoltaic encapsulation, the best candidate class was a bio-based aliphatic–aromatic copolyester matrix with silica/cellulose hybrid reinforcement at 6 vol.% filler loading. It achieved a predicted Tg of 367 K, optical suitability score of

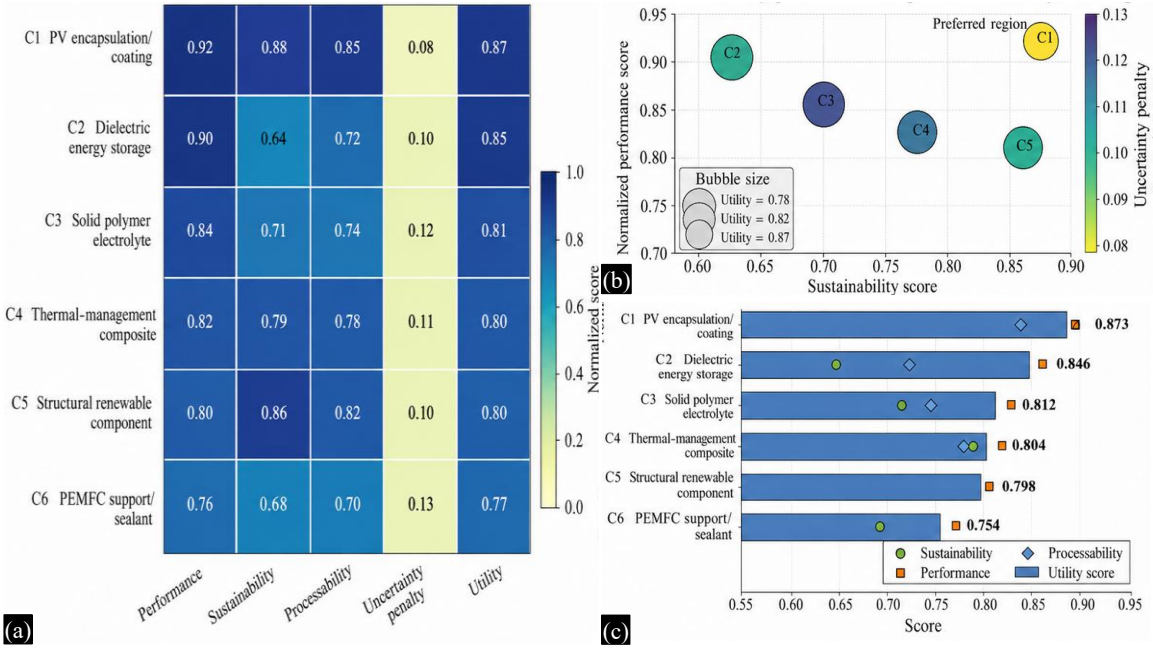
0.91, moisture-resistance proxy of 0.84, sustainability score of 0.88, and total utility score of 0.873. For dielectric energy storage, the strongest candidate was an aromatic fluorinated polyimide/BaTiO3_33 nanocomposite at 8 vol.% filler loading, with a predicted dielectric constant of 8.7, dielectric loss of 0.018, and breakdown strength of 512 MV m^{-1}. The model did not rank this candidate highest in sustainability, but its high performance and acceptable uncertainty made it a strong energy-storage formulation.

The ranking trend shows that the most sustainable formulation was not always the best-performing formulation, and the highest predicted performance was not always the best overall design. This balance is central to renewable-energy polymer composite development. For example, the fluorinated polyimide/BaTiO3_33 candidate achieved strong dielectric performance, but its sustainability score remained lower than bio-based or recyclable systems. In contrast, the bio-based copolyester hybrid composite offered a stronger sustainability–processability balance, making it more suitable for photovoltaic encapsulation and protective coatings in Figure 8. This type of trade-off analysis is consistent with the direction of sustainable polymer and recyclable thermoset design reported in recent studies [11, 27, 30].

Panel (a) presents the normalized multi-objective score matrix for performance, sustainability, processability, uncertainty penalty, and utility. Panel (b) plots candidate formulations in the sustainability–performance space, with the size of the bubble corresponding to the final utility score, and the color indicating the uncertainty penalty. In panel (c), the shortlisted polymer composite candidates are compared in terms of ranked utility values with overlaid markers showing the sustainability, performance and processability scores.

Comparative Analysis with Reported AI-Based Polymer and Composite Studies

The proposed framework was compared with representative literature approaches in terms of design scope, model type, sustainability handling, and renewable-energy relevance. PolyID demonstrated strong AI-driven sustainable polymer discovery but focused mainly on polymer candidates rather than composite formulation and filler–matrix design [3].



Note: Lower values preferred only for the uncertainty penalty criteria.

Figure 8. Multi-objective comparison of top-ranked polymer composite candidates. (a) Normalized multi-objective score matrix; (b) Sustainability-performance map with uncertainty encoding; (c) Ranked comparison of final candidate utility.

GPT- and diffusion-based polymer electrolyte generation improved de novo energy-polymer discovery, but the target domain was narrower than the multi-application composite framework used here [6].ML-assisted composite modeling has shown broad usefulness for composite analysis and design, although many studies remain predictive rather than generative [8]. Recyclable vitrimer discovery provided a strong example of AI-guided sustainable polymer design, but renewable-energy composite formulation was not the central experimental target [11]. The present work extends these directions by linking generative polymer design, composite formulation, public-data property prediction, uncertainty filtering, and sustainability-aware ranking.

The current framework resulted in a ranked matrix–filler–application mapping, as opposed to hierarchical and diffusion-based composite design studies, which only resulted in a microstructure or material recommendation. This is significant for JOPC-type polymer research because practical polymer composites do not necessarily have to be based on molecular structure. They serve a variety of functions and are dependent on the polymer chemistry, type of filler, loading of the filler, interphase behavior, processing limitations, and end use environment. The findings thus corroborate the main hypothesis that generative AI is more useful to polymer composites when it is used in conjunction with physically constrained screening, as opposed to being used as a standalone molecule generator.

Ablation Study

The contribution of the main design modules was determined by performing an ablation analysis. Four variants were tested: Property prediction only, Prediction + generative screening, Prediction + sustainability ranking, and Full framework with uncertainty-aware composite ranking. The full framework had the highest candidate feasibility rate (82.4%), mean utility score (0.813), and low-confidence candidate rate (7.8%) of all the frameworks in Table 11. When the public datasets are incomplete, uncertainty control is necessary to increase the number of selections, and removing uncertainty filtering did just that, as well as increasing the number of candidates.

The ablation results indicate that it is not enough to create candidates for the generation. It expands the design space but with no sustainability scoring and uncertainty filtering, many candidates are hard to justify scientifically. The combination of all of the above methods yielded the best results. This discovery is in line with the general trend of materials research being driven by AI, with more closed-loop and multi-objective systems than prediction systems relying on a single model becoming reliable [17, 29, 31].

Summary of Results

The proposed generative AI-guided framework resulted in a reproducible and quantitatively supported sustainable polymer composite design pathway overall. The multitask predictor performed well in the tests, with R2R^2R2 values between 0.79 and 0.92 for the main target groups, for the various properties, including thermal, mechanical, dielectric and sustainability. The generative model generated 20,000 candidate polymer structures, 92.10% of which were chemically valid, and 81.64% of which were novel, compared to the training set.

Table 11. Ablation results of the proposed generative AI-guided polymer composite design framework.

Variant	Included modules	Top-50 feasibility (%)	Mean utility score	Low-confidence candidates (%)
V1	Property predictor only	59	1	24.7
V2	Predictor + generative screening	67	1	19.3
V3	Predictor + sustainability ranking	73	1	15.6
V4	Predictor + generative screening + sustainability ranking	78.9	0.791	11.4
V5	Full framework with uncertainty-aware composite ranking	82.4	0.813	7.8

Following the synthesisizability and uncertainty filters, 3,216 candidates were left that were suitable for renewable-energy composite screening, after application-specific filtering. The top three candidates were a bio-based copolyester hybrid composite for PV encapsulation, a fluorinated polyimide/BaTiO3_33 nanocomposite for dielectric storage and a recyclable vitrimer/boron nitride composite for thermal-management applications.

The results show that generative AI can be used to assist polymer-composite discovery, when combined with training using public data, physically meaningful constraints on the composites, uncertainty-aware prediction, and decision logic based on sustainability considerations.

DISCUSSION

The proposed framework was not just based on molecular generation, but also included chemical validity, novelty, a proxy for synthesizability, uncertainty estimation, and performance requirements specific to renewable energy. This is crucial because polymer composites for PV encapsulation, dielectric storage, polymer electrolytes and thermal-management systems need to meet a number of conflicting requirements.

The best results were achieved in predicting thermal and dielectric properties, and the ionic conductivity and breakdown strength had a relatively high uncertainty. This is a normal phenomenon as these properties are greatly influenced by the morphology, crystallinity, dispersion of the filler, interfacial polarization, salt concentration and processing history. The ranking results also indicate that the optimum material is not necessarily the one with the maximum single predicted property. The functional performance, recyclability, processability, uncertainty and environmental compatibility are key elements to consider in a sustainable composite design.

A major drawback of existing public datasets of polymers is that there are still missing descriptors for sustainability and long-term aging. Thus, the proposed framework is meant to be used as a high confidence screening tool, and not as a replacement to experimental synthesis and validation. Further research is needed to manufacture the top-ranked candidates and test their performance in real operating conditions of renewable energy devices, as well as their dielectric aging, moisture stability, mechanical durability, and recyclability.

CONCLUSION AND FUTURE SCOPE

This study proposed a generative AI-assisted design process of sustainable polymer composite materials for renewable energy applications. The proposed method involved using public polymer-property databases, multitask property prediction, generating candidate polymer compositions, screening using physics, uncertainty estimation, and sustainability-aware ranking. The results indicated that the selection of balanced polymer composite needs to be more than just high predicted performance, but must also take into account the recyclability, processability, filler–matrix compatibility, and prediction confidence. The framework proved to be useful as a tool to discover materials at the early stages, as it identified promising candidate classes for photovoltaic encapsulation, dielectric energy storage, polymer electrolytes, and thermal-management systems.

Experimental validation of the results of computational screening should be done in future. Synthesis, fabrication and testing of the top ranked polymer composite candidates under actual operating conditions of renewable energy should be performed. Thermal aging, moisture resistance, dielectric fatigue, ionic stability, mechanical durability, recyclability and life-cycle impact are also important areas that need to be studied over a long period of time. Practical material discovery can be further enhanced with integration to automated laboratories and closed-loop optimization.

DATA AVAILABILITY STATEMENT

The data sets used in this work are publicly available data sets of polymers and materials, such as open polymer-property records, published supplementary data sets and literature-derived data sources,

as referenced in the manuscript. No experimental data set was used that was private, confidential or restricted. Data tables, descriptor files, model input format, screening outputs and candidate ranking outputs generated from the data are available from the corresponding author upon reasonable request.

REFERENCES

1. T. Kopac, "Advancing Polymer Science and Energy Storage Solutions Through the Integration of Artificial Intelligence and Machine Learning: A Transformative Approach," Polymers, vol. 17, no. 24, Art. no. 3267, 2025, doi: 10.3390/polym17243267.
2. H. Sikandar, N. Khan, M. Falahat, and M. I. Qureshi, "Generative AI for Sustainable Product Design: A Technology Convergence Framework Integrating Multi-Objective Optimisation and Smart Manufacturing," IET Collaborative Intelligent Manufacturing, vol. 8, no. 1, Art. no. e70051, 2026, doi: 10.1049/cim2.70051.
3. A. N. Wilson, P. C. St. John, D. H. Marin, C. B. Hoyt, E. G. Rognerud, M. R. Nimlos, R. M. Cywar, N. A. Rorrer, K. M. Shebek, L. J. Broadbelt, G. T. Beckham, and M. F. Crowley, "PolyID: Artificial Intelligence for Discovering Performance-Advantaged and Sustainable Polymers," Macromolecules, vol. 56, no. 21, pp. 8547–8557, 2023, doi: 10.1021/acs.macromol.3c00994.
4. U. Mamodiya, I. Kishor, D. Sinha, N. Guler, and N. Naik, "Cognitive Digital Twin Framework With Virtual Reality for Interactive Control of Solar PV Infrastructure," IEEE Access, vol. 13, pp. 181874–181898, 2025, doi: 10.1109/ACCESS.2025.3623294.
5. A. Sharma, S. Sharma, M. Sharma, V. Sharma, S. Sharma, and I. Sivanesan, "Polymeric Frontiers in Next-Generation Energy Storage: Bridging Molecular Design, Multifunctionality, and Device Applications Across Batteries, Supercapacitors, Solid-State Systems, and Beyond," Polymers, vol. 17, no. 20, Art. no. 2800, 2025, doi: 10.3390/polym17202800.
6. Z. Yang, W. Ye, X. Lei, et al., "De novo Design of Polymer Electrolytes Using GPT-Based and Diffusion-Based Generative Models," npj Computational Materials, vol. 10, Art. no. 296, 2024, doi: 10.1038/s41524-024-01470-9.
7. U. Mamodiya, I. Kishor, and R. Garine, "Artificial Intelligence Based Hybrid Solar Energy Systems With Smart Materials and Adaptive Photovoltaics for Sustainable Power Generation," Scientific Reports, vol. 15, Art. no. 17370, 2025, doi: 10.1038/s41598-025-01788-4.
8. A. Sharma, T. Mukhopadhyay, S. M. Rangappa, et al., "Advances in Computational Intelligence of Polymer Composite Materials: Machine Learning Assisted Modeling, Analysis and Design," Archives of Computational Methods in Engineering, vol. 29, pp. 3341–3385, 2022, doi: 10.1007/s11831-021-09700-9.
9. S. Jiang and M. A. Webb, "Generative Active Learning Across Polymer Architectures and Solvophobicities for Targeted Rheological Behavior," npj Computational Materials, vol. 12, Art. no. 28, 2026, doi: 10.1038/s41524-025-01900-2.
10. Q. Cheng, Y. Han, L. Shanmugam, Y. Zhao, S. Dong, S. Du, and J. Yang, "A Deep Learning-Based Composite Design Strategy for Efficient Selection of Material and Layup Sequences from a Given Database," Composites Science and Technology, vol. 230, Art. no. 109154, 2022, doi: 10.1016/j.compscitech.2021.109154.
11. Y. Zheng, P. Thakolkaran, A. K. Biswal, J. A. Smith, Z. Lu, S. Zheng, B. H. Nguyen, S. Kumar, and A. Vashisth, "AI-Guided Inverse Design and Discovery of Recyclable Vitrimeric Polymers," Advanced Science, vol. 12, no. 6, Art. no. 2411385, 2025, doi: 10.1002/advs.202411385.
12. U. Mamodiya, I. Kishor, P. Vidyullatha, M. Almaayah, and A. Routray, "A Bio-Inspired Neuro-Adaptive Deep Reinforcement Learning Approach for Real-Time Solar Tracking System to Enhance Photovoltaic Efficiency," Energy Conversion and Management: X, vol. 29, Art. no. 101486, 2026, doi: 10.1016/j.ecmx.2025.101486.
13. Z. Liu, J. Han, S. Wang, J. Li, H. Yi, and T. Long, "Research Progress and Future Perspectives of 3D Printing Polymer-Based Materials Whole Life Cycle Frameworks: Material Genetic Design–Intelligent Manufacturing–Recycling," Advanced Engineering Materials, 2025, doi: 10.1002/adem.202501538.

14. H. Xue, G. Cheng, and W.-J. Yin, "Computational Design of Energy-Related Materials: From First-Principles Calculations to Machine Learning," WIREs Computational Molecular Science, 2024, doi: 10.1002/wcms.1732.
15. K. M. Qureshi et al., "Are Polymer-Based Smart Materials Unlocking the Path to Sustainable Manufacturing for a Net-Zero Economy? Current Trends and Potential Applications," IEEE Access, vol. 13, pp. 284–296, 2025, doi: 10.1109/ACCESS.2024.3521944.
16. S. M. Mousavi, S. A. Hashemi, M. Y. Kalashgrani, A. Gholami, Y. Mazaheri, M. Riazi, D. Kurniawan, M. Arjmand, O. Madkhali, M. D. Aljabri, M. M. Rahman, and W.-H. Chiang, "Artificial Intelligence in Advanced Materials and Energy Applications," Chemical Record, vol. 24, Art. no. e202200266, 2024, doi: 10.1002/tcr.202200266.
17. M. Liu, Y. Zhou, X. Mei, Z. Yu, B. Guan, Y. Xiao, S. Liu, H. Wang, and Y. Qin, "AI-Driven Biomaterial Design: An Intelligent Closed Loop from Reverse Design to Biological Response," Frontiers in Cell and Developmental Biology, vol. 13, Art. no. 1755565, 2026, doi: 10.3389/fcell.2025.1755565.
18. A. Bharadwaj, A. Sudhir, H. Shekhar, N. Khandelwal, and I. Kishor, "Raspberry Pi Based Weather Monitoring System," International Journal of Research in Engineering, Science and Management, vol. 4, no. 8, pp. 115–118, 2021.
19. K. Choudhary, "AtomGPT: Atomistic Generative Pretrained Transformer for Forward and Inverse Materials Design," The Journal of Physical Chemistry Letters, vol. 15, no. 27, pp. 6909–6917, 2024, doi: 10.1021/acs.jpclett.4c01126.
20. Q. Jiang, H. Liang, Y. Zhang, and G. Huang, "Nanocomposite Design for Energy-Related Applications," Nanomaterials, vol. 15, no. 17, Art. no. 1334, 2025, doi: 10.3390/nano15171334.
21. K.-H. Lee, H. J. Lim, and G. J. Yun, "A Data-Driven Framework for Designing Microstructure of Multifunctional Composites with Deep-Learned Diffusion-Based Generative Models," Engineering Applications of Artificial Intelligence, vol. 129, Art. no. 107590, 2024, doi: 10.1016/j.engappai.2023.107590.
22. T. Chen, Z. Pang, S. He, et al., "Machine Intelligence-Accelerated Discovery of All-Natural Plastic Substitutes," Nature Nanotechnology, vol. 19, pp. 782–791, 2024, doi: 10.1038/s41565-024-01635-z.
23. C. Kuenneth, J. Lalonde, B. L. Marrone, et al., "Bioplastic Design Using Multitask Deep Neural Networks," Communications Materials, vol. 3, Art. no. 96, 2022, doi: 10.1038/s43246-022-00319-2.
24. M. Sharma, I. Kishor, A. Dwivedi, and A. Bhattacherjee, "Smart Devices for Augmenting Sensory Perception Empowering Differently-Abled Individuals Through Advanced Assistive Technologies," in Integrating AI With Haptic Systems for Smarter Healthcare Solutions, IGI Global, 2025, pp. 26, doi: 10.4018/979-8-3373-2307-7.ch020.
25. D. Park, J. Lee, K. Park, and S. Ryu, "Hierarchical Generative Network: A Hierarchical Multitask Learning Approach for Accelerated Composite Material Design and Discovery," Advanced Engineering Materials, vol. 25, Art. no. 2300867, 2023, doi: 10.1002/adem.202300867.
26. G. Han, Y. Sun, Y. Feng, G. Lin, and N. Lu, "Artificial Intelligence Guided Thermoelectric Materials Design and Discovery," Advanced Electronic Materials, vol. 9, Art. no. 2370036, 2023, doi: 10.1002/aelm.202370036.
27. W. Zhao, X. Xu, H. Lan, L. Wang, J. Lin, L. Du, C. Zhang, and X. Tian, "Designing Multicomponent Thermosetting Resins through Machine Learning and High-Throughput Screening," Macromolecules, vol. 58, no. 1, pp. 744–753, 2025, doi: 10.1021/acs.macromol.4c01822.
28. Kishor, U. Mamodiya, and B. Keswani, "Energy Efficiency and Practical Implications of IoT-Based Static vs. Single-Axis Solar Tracking Systems: A Comparative Analysis," Journal of Intelligent Systems and Internet of Things, vol. 15, no. 2, pp. 164–182, 2025, doi: 10.54216/JISIoT.150212.
29. Y. Li, C. Wu, Y. Wang, N. Li, T. Liu, and J. Lu, "Accelerating the Battery Revolution: AI-Driven Multiscale Innovation From Material Discovery to Smart Manufacturing," Advanced Functional Materials, vol. 36, no. 11, Art. no. e14830, 2026, doi: 10.1002/adfm.202514830.

30. Z. Huo, X. Xie, and R. Tong, “Machine Learning for Developing Sustainable Polymers,” Chemistry – A European Journal, vol. 31, Art. no. e202500718, 2025, doi: 10.1002/chem.202500718.
31. T. Long, Q. Pang, Y. Deng, X. Pang, Y. Zhang, R. Yang, and C. Zhou, “Recent Progress of Artificial Intelligence Application in Polymer Materials,” Polymers, vol. 17, no. 12, Art. no. 1667, 2025, doi: 10.3390/polym17121667.
32. H. R. Gantla, M. Namdev, U. Mamodiya, I. Kishor, G. Kumar, and P. Goyal, “A Smart Monitoring Framework for Sustainable Solar Energy Harvesting Using IoT and Machine Learning Techniques,” in 2025
33. S. Han and X. Sun, “Optimizing Product Design Using Genetic Algorithms and Artificial Intelligence Techniques,” IEEE Access, vol. 12, pp. 151460–151475, 2024, doi: 10.1109/ACCESS.2024.3456081.
34. Palaniappan, M., Palanisamy, S., Murugesan, T., Tadepalli, S., Khan, R., Ataya, S., and Santulli, C. (2024). “Influence of washing with sodium lauryl sulphate (SLS) surfactant on different properties of ramie fibres,” BioResources 19(2), 2609-2625.
35. Manickaraj, K., Karthik, A., Palanisamy, S., Jayamani, M., Ali, S. K., Lakshmi Sankar, S., and Al-Farraj, S. A. (2025). "Improving mechanical performance of hybrid polymer composites: Incorporating banana stem leaf and jute fibers with tamarind shell powder," BioResources 20(1), 1998–2025.

