TROY TECHNICAL

MATERIALS INFORMATICS
WEEKLY REPORT

October 24 - October 30, 2026 | Issue 2026-W43

Closed-loop learning from data quality to experimentally validated material performance

13	10
ANALYZED ARTICLES	PRIORITY THEMES

Issued 2026-10-30 | English edition | fixed approved cover artwork



Materials Informatics: the boundary moved from isolated claims to qualification evidence

13 English articles were reviewed independently from the Japanese edition. Ten themes are retained for management decisions.

SIGNAL 01 | DATA CAPTURE

MARVEL-AI: A*STAR's AI-Driven Materials Discovery Framework

A*STAR has launched MARVEL-AI to accelerate the transition of advanced materials from discovery to industrial deployment using an AI-enabled closed-loop system.

SIGNAL 02 | MODELING

Graph Neural Networks: Accelerating Electrocatalyst Design for 2026

Graph neural networks (GNNs) are replacing traditional trial-and-error and DFT screening to accelerate the discovery of high-performance electrocatalysts for green hydrogen production.

SIGNAL 03 | CANDIDATE DESIGN

Open Polymer Datasets: Reuse Quality Audit across 8 Repositories (2020-2025)

A cross-repository audit of 53,038 open polymer datasets reveals that the choice of repository is the primary determinant of data reusability, outweighing science or publication year.

SIGNAL 04 | EXPERIMENT

Lithium Metal Batteries: AI-Driven Discovery of 368 Non-Fluorinated Diluents

Researchers at Seoul National University developed an AI algorithm that screened 1.32 million molecules to identify 368 promising non-fluorinated electrolyte diluents for lithium metal batteries.

DECISION

Focus this week's management attention on closed-loop learning from data quality to experimentally validated material performance. Move only when the linked evidence can be reproduced under your own operating conditions.

THREE MANAGEMENT MOVES

1. Buyers: select high-value decisions where data lineage and experimental feedback can be measured.
2. Suppliers: provide machine-readable composition, process, test conditions, uncertainty and failure data.
3. Executives: track dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EVIDENCE DISCIPLINE

FACT states what the collected English source reports. BUSINESS READ and ACTION are editorial interpretations. UNKNOWN lists information that the source file does not establish. A linked article is not treated as independent confirmation of every claim.

Five numbers or evidence anchors that require conditions

143%
A10

SOURCE-BOUND CONDITION
While deposit volumes grew 143% between 2020 and 2025, average documentation quality remained essentially flat, creating a widening gap between data volume and reuse value.

1.32 million
A13

SOURCE-BOUND CONDITION
Researchers at Seoul National University developed an AI algorithm that screened 1.32 million molecules to identify 368 promising non-fluorinated electrolyte diluents for lithium metal batteries.

P03
A10

SOURCE-BOUND CONDITION
A cross-repository audit of 53,038 open polymer datasets reveals that the choice of repository is the primary determinant of data reusability, outweighing science or publication year.

P04
A13

SOURCE-BOUND CONDITION
Researchers at Seoul National University developed an AI algorithm that screened 1.32 million molecules to identify 368 promising non-fluorinated electrolyte diluents for lithium metal batteries.

P05
A02

SOURCE-BOUND CONDITION
Researchers from Columbia Engineering and the University of Cambridge have developed a benchmark to determine if machine learning models for materials science derive macroscopic properties from correct underlying.

HOW TO USE THESE NUMBERS
Each value is shown with the condition stated in the collected English article. Do not compare values across different test methods, device sizes, operating temperatures or commercial stages without normalizing those conditions.

Where the value chain moved

Each stage combines one leading theme and one supporting theme. The report treats handoffs as evidence transfers, not decorative arrows.

1

Data capture

P01/P06

A*STAR has launched MARVEL-AI to accelerate the transition of advanced materials from discovery to industrial deployment.

2

Modeling

P02/P07

Graph neural networks (GNNs) are replacing traditional trial-and-error and DFT screening to accelerate the discovery of.

3

Candidate design

P03/P08

A cross-repository audit of 53,038 open polymer datasets reveals that the choice of repository is the primary determinant of.

4

Experiment

P04/P09

Researchers at Seoul National University developed an AI algorithm that screened 1.32 million molecules to identify 368.

5

Scale & governance

P05/P10

Researchers from Columbia Engineering and the University of Cambridge have developed a benchmark to determine if machine.

EVIDENCE REQUIRED AT EACH HANDOFF

Data capture to Modeling	Transfer test conditions, acceptance criteria, failure data, owner and decision date. Recheck dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.
Modeling to Candidate design	Transfer test conditions, acceptance criteria, failure data, owner and decision date. Recheck dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.
Candidate design to Experiment	Transfer test conditions, acceptance criteria, failure data, owner and decision date. Recheck dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.
Experiment to Scale & governance	Transfer test conditions, acceptance criteria, failure data, owner and decision date. Recheck dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

COMMON CONTROL POINT

No theme moves to a commercial commitment until a named owner has reproduced the relevant evidence and documented dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

Priority map: maturity and next decision locus



P01	MARVEL-AI: A*STAR's AI-Driven Materials Discovery Framework System / adoption | cue: launched
P02	Graph Neural Networks: Accelerating Electrocatalyst Design for 2026 System / adoption | cue: trial
P03	Open Polymer Datasets: Reuse Quality Audit across 8 Repositories (2020-2025) Tech / process | cue: no system/business cue
P04	Lithium Metal Batteries: AI-Driven Discovery of 368 Non-Fluorinated Diluents System / adoption | cue: trial
P05	AI Materials Models: New Benchmark for Physics-Aware Accuracy System / adoption | cue: adoption
P06	qNEP Potential: Sodium Battery SEI Formation and Simulation Specs Tech / process | cue: no system/business cue
P07	Epoxy Degradation Analysis: PCA-Based Spectral Fingerprinting Tech / process | cue: no system/business cue
P08	CLaC Model: MIT's AI for Crystal Structure and Text Retrieval Tech / process | cue: no system/business cue
P09	MXene Discovery: AI-Driven Identification of 33 2D Materials Tech / process | cue: no system/business cue
P10	Transition Metal Carbides: Composition-Based Melting Point Prediction Tech / process | cue: no system/business cue

REPRODUCIBLE POSITIONING RULE

Horizontal position uses five discrete stages: Research, Prototype, Joint evaluation, Product adoption and Scale. The vertical bands classify the next decision locus: Tech / process, System / adoption, or Business / rules (including supply systems). Article-title cues take precedence; the overview is used only when the title is silent. These are categories, not impact ranks. Bubble diameter reflects the editorial score.

FACT is source-bound. BUSINESS READ, ACTION and UNKNOWN are explicit editorial layers.

P01 MARVEL-AI: A*STAR's AI-Driven Materials Discovery Framework

A07 | Data capture | Product adoption

Date not stated

FACT

A*STAR has launched MARVEL-AI to accelerate the transition of advanced materials from discovery to industrial deployment using an AI-enabled closed-loop system. The initiative demonstrates significant throughput gains, such as a 25-fold increase in polymer sample production and 400-fold increase in alloy screening. This framework aims to bridge the gap between laboratory research and real-world application across semiconductors, energy, and chemicals.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P01.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 21/25 | MATURITY 4/5 | DECISION LOCUS: SYSTEM / ADOPTION

P02 Graph Neural Networks: Accelerating Electrocatalyst Design for 2026

A09 | Modeling | Product adoption

Date not stated

FACT

Graph neural networks (GNNs) are replacing traditional trial-and-error and DFT screening to accelerate the discovery of high-performance electrocatalysts for green hydrogen production. By learning material properties directly from atomic structures, GNNs bypass manual feature engineering to identify low-cost alternatives to platinum and iridium.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P02.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 20/25 | MATURITY 4/5 | DECISION LOCUS: SYSTEM / ADOPTION

FACT is source-bound. BUSINESS READ, ACTION and UNKNOWN are explicit editorial layers.

P03 Open Polymer Datasets: Reuse Quality Audit across 8 Repositories (2020-2025)

A10 | Candidate design | Research

Date not stated

FACT

A cross-repository audit of 53,038 open polymer datasets reveals that the choice of repository is the primary determinant of data reusability, outweighing science or publication year. While deposit volumes grew 143% between 2020 and 2025, average documentation quality remained essentially flat, creating a widening gap between data volume and reuse value.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P03.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 20/25 | MATURITY 1/5 | DECISION LOCUS: TECH / PROCESS

P04 Lithium Metal Batteries: AI-Driven Discovery of 368 Non-Fluorinated Diluents

A13 | Experiment | Joint evaluation

Date not stated

FACT

Researchers at Seoul National University developed an AI algorithm that screened 1.32 million molecules to identify 368 promising non-fluorinated electrolyte diluents for lithium metal batteries. Three of these candidates were experimentally verified in pouch cells, demonstrating stable charge-discharge performance. This approach significantly reduces the time and cost of material discovery by replacing trial-and-error synthesis with physics-based machine learning.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P04.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 20/25 | MATURITY 3/5 | DECISION LOCUS: SYSTEM / ADOPTION

FACT is source-bound. BUSINESS READ, ACTION and UNKNOWN are explicit editorial layers.

P05 AI Materials Models: New Benchmark for Physics-Aware Accuracy

A02 | Scale & governance | Pilot

Date not stated

FACT

Researchers from Columbia Engineering and the University of Cambridge have developed a benchmark to determine if machine learning models for materials science derive macroscopic properties from correct underlying physics. The study reveals that models can predict thermal conductivity accurately while remaining wrong about the atomic vibrations that cause it, highlighting a critical reliability gap for industrial adoption.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P05.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 20/25 | MATURITY 2/5 | DECISION LOCUS: SYSTEM / ADOPTION

P06 qNEP Potential: Sodium Battery SEI Formation and Simulation Specs

A12 | Data capture | Pilot

Date not stated

FACT

Researchers have developed a charge-aware neuroevolution potential (qNEP) that extends simulation timescales for sodium battery interfaces to 100 nanoseconds. This machine-learning approach overcomes previous bottlenecks in observing the formation of the solid electrolyte interphase (SEI), enabling the study of nanosecond-scale densification and phase separation.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P06.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 20/25 | MATURITY 2/5 | DECISION LOCUS: TECH / PROCESS

FACT is source-bound. BUSINESS READ, ACTION and UNKNOWN are explicit editorial layers.

P07 Epoxy Degradation Analysis: PCA-Based Spectral Fingerprinting

A06 | Modeling | Pilot

Date not stated

FACT

Researchers used principal component analysis (PCA) to link the spectral fingerprints of degrading amine-cured epoxy adhesives to the loss of mechanical strength. The study reveals that degradation behavior varies significantly depending on the curing agent used, with some systems showing non-monotonic strength changes.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P07.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 20/25 | MATURITY 2/5 | DECISION LOCUS: TECH / PROCESS

P08 CLaC Model: MIT's AI for Crystal Structure and Text Retrieval

A08 | Candidate design | Pilot

Date not stated

FACT

MIT researchers developed CLaC, a vision-language model that aligns 3D crystal structures with natural language descriptions using contrastive learning. By utilizing synthetic narratives to overcome data bias, the model achieved a 95.4 percent accuracy rate in matching structures to compositions in zero-shot tests. This method provides a scalable way to retrieve promising materials for applications such as solid-state battery electrolytes.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P08.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 19/25 | MATURITY 2/5 | DECISION LOCUS: TECH / PROCESS

FACT is source-bound. BUSINESS READ, ACTION and UNKNOWN are explicit editorial layers.

P09 MXene Discovery: AI-Driven Identification of 33 2D Materials

A01 | Experiment | Research

Date not stated

FACT

An international research team led by Linköping University identified 33 previously overlooked MXene candidates by applying machine learning to decades-old crystal data. The team successfully synthesized 18 multilayer phases, including 11 new rare-earth compositions, utilizing a high-temperature synthesis method. This workflow demonstrates how AI can rediscover buried materials from legacy databases to accelerate 2D material development.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P09.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 19/25 | MATURITY 1/5 | DECISION LOCUS: TECH / PROCESS

P10 Transition Metal Carbides: Composition-Based Melting Point Prediction

A03 | Scale & governance | Research

Date not stated

FACT

Machine learning analysis reveals that elemental composition alone is sufficient to predict the melting temperatures of transition metal carbides. Adding complex descriptors like formation energies and bond geometry produced no statistically significant gain in accuracy, simplifying the data requirements for predicting ultra-high-temperature ceramic properties.

WHY IT MATTERS

The decision relevance is closed-loop learning from data quality to experimentally validated material performance. The signal should be tested at the application and operating-system level, not accepted from a headline alone.

BUSINESS READ

For operating companies, compare the reported change with current qualification, sourcing and product-roadmap assumptions. For suppliers, provide machine-readable composition, process, test conditions, uncertainty and failure data.

ACTION

Within 30 days, select high-value decisions where data lineage and experimental feedback can be measured; record the evidence and owner against P10.

UNKNOWN

Still unproven or incompletely stated: dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.

EDITORIAL SCORE 19/25 | MATURITY 1/5 | DECISION LOCUS: TECH / PROCESS

Playbook and 0-5 year roadmap

STAKEHOLDER	NOW | 0-30 DAYS	NEXT | 1-12 MONTHS	LATER | 1-5 YEARS
Operating companies	select high-value decisions where data lineage and experimental feedback can be measured	Run production-like qualification with explicit acceptance and stop criteria.	Build a portfolio and architecture that can absorb technology and supplier changes.
Materials, equipment and technology suppliers	provide machine-readable composition, process, test conditions, uncertainty and failure data	Create shared reliability, process-window and failure-analysis data with customers.	Standardize interfaces, traceability, lifecycle support and recovery services.
Trading, investment and legal teams	Map dataset bias, reproducibility, IP, model transfer, experiment cost and adoption.	Contract for capacity, quality, change notice, liability, alternatives and exit.	Diversify regional, technical and customer concentration risk.

ROADMAP

NOW | 0-30 DAYS
Evidence ledger
Owner: Strategy + quality
Deliverable: claim, source, condition, owner

NOW | 0-30 DAYS
Risk screen
Owner: Procurement + legal
Deliverable: unknowns, alternatives, stop

NEXT | 1-12 MONTHS
Joint qualification
Owner: R&D + supplier
Deliverable: process window and failure data

NEXT | 1-12 MONTHS
Commercial gate
Owner: Business + finance
Deliverable: unit economics and capacity

LATER | 1-5 YEARS
Platform strategy
Owner: Executive team
Deliverable: portfolio and interface roadmap

LATER | 1-5 YEARS
Service layer
Owner: Operations + channel
Deliverable: monitoring, maintenance and recovery

GO / NO-GO GATES

GATE	OWNER	REQUIRED EVIDENCE	NO-GO CONDITION
G1 Evidence	**Quality**	Source, test conditions and reproducibility	No traceable evidence
G2 Integration	**R&D**	Interface, process window and failure analysis	Cannot meet system conditions
G3 Commercial	**Business**	Capacity, total cost, contract and alternatives	Economics or supply cannot be supported
G4 Lifecycle	**Operations**	Monitoring, maintenance, change notice and exit	No accountable operating plan

ANALYZED ARTICLES | ITEMS 01-13

Every title and URL is displayed in full. URLs wrap without shortening and remain clickable. A verified-reference label means the translated HTML did not retain its original source URL.

ID	FULL ARTICLE TITLE	FULL SOURCE / VERIFIED REFERENCE URL - CLICKABLE
A01	MXene Discovery: AI-Driven Identification of 33 2D Materials	https://xenospectrum.com/en/ai-rediscovered-multilayer-mxenes-science-2026/
A02	AI Materials Models: New Benchmark for Physics-Aware Accuracy	https://bioengineer.org/new-benchmark-tests-whether-ai-materials-models-get-physics-right/
A03	Transition Metal Carbides: Composition-Based Melting Point Prediction	https://bioengineer.org/machine-learning-shows-composition-alone-predicts-melting-in-transition-metal-carbides/
A04	EdenGNN: Equivariant Neural Network for Faster Charge Density Prediction	https://bioengineer.org/equivariant-neural-network-predicts-charge-density-over-an-order-of-magnitude-faster/
A05	NVIDIA Tensor Cores: Speedups for ML Interatomic Potentials	https://blockchain.news/news/nvidia-blog-tensor-core-speedups-ml-interatomic-potentials
A06	Epoxy Degradation Analysis: PCA-Based Spectral Fingerprinting	https://bioengineer.org/research-report-unsupervised-learning-based-analysis-of-degradation-behavior-in-filler-containing-amine-cured-epoxy-formulations/
A07	MARVEL-AI: A*STAR's AI-Driven Materials Discovery Framework	https://www.a-star.edu.sg/news/press-releases/marvel-ai-accelerating-advanced-materials-innovation-from-discovery-to-deployment
A08	CLaC Model: MIT's AI for Crystal Structure and Text Retrieval	https://www.intelligentliving.co/mit-taught-ai-read-crystals-find-better/
A09	Graph Neural Networks: Accelerating Electrocatalyst Design for 2026	https://bioengineer.org/graph-neural-networks-reshape-electrocatalyst-design-for-green-hydrogen/
A10	Open Polymer Datasets: Reuse Quality Audit across 8 Repositories (2020-2025)	https://bioengineer.org/reuse-quality-of-open-polymer-datasets-a-reproducible-cross-repository-audit-2020-2025/
A11	OrgAIcat: ML model for asymmetric organocatalysis selectivity	https://bioengineer.org/machine-learning-model-turns-decades-of-literature-data-into-a-fast-route-to-highly-selective-organocatalysis/
A12	qNEP Potential: Sodium Battery SEI Formation and Simulation Specs	https://bioengineer.org/machine-learning-simulations-reveal-how-sodium-batteries-form-their-protective-interface/
A13	Lithium Metal Batteries: AI-Driven Discovery of 368 Non-Fluorinated Diluents	https://www.aitimes.kr/news/articleView.html?idxno=42475

MaterialsInformatics — Selected Articles

Date: 2026-10-31

Articles: 13

Table of Contents

#01 MXene Discovery: AI-Driven Identification of 33 2D Materials

#02 AI Materials Models: New Benchmark for Physics-Aware Accuracy

#03 Transition Metal Carbides: Composition-Based Melting Point Prediction

#04 EdenGNN: Equivariant Neural Network for Faster Charge Density Prediction

#05 NVIDIA Tensor Cores: Speedups for ML Interatomic Potentials

#06 Epoxy Degradation Analysis: PCA-Based Spectral Fingerprinting

#07 MARVEL-AI: A*STAR's AI-Driven Materials Discovery Framework

#08 CLaC Model: MIT's AI for Crystal Structure and Text Retrieval

#09 Graph Neural Networks: Accelerating Electrocatalyst Design for 2026

#10 Open Polymer Datasets: Reuse Quality Audit across 8 Repositories (2020–2025)

#11 OrgAIcat: ML model for asymmetric organocatalysis selectivity

#12 qNEP Potential: Sodium Battery SEI Formation and Simulation Specs

#13 Lithium Metal Batteries: AI-Driven Discovery of 368 Non-Fluorinated Diluents

#01 MXene Discovery: AI-Driven Identification of 33 2D Materials

Published October 30, 2026 xenospectrum

OVERVIEW

An international research team led by Linköping University identified 33 previously overlooked MXene candidates by applying machine learning to decades-old crystal data. The team successfully synthesized 18 multilayer phases, including 11 new rare-earth compositions, utilizing a high-temperature synthesis method. This workflow demonstrates how AI can rediscover buried materials from legacy databases to accelerate 2D material development.

IN DEPTH

Key Findings

Researchers have identified 33 candidates for MXenes—two-dimensional materials—that had been synthesized in the 1970s but remained unrecognized as such until now. By combining machine-learning analysis of crystal structure databases with a review of past literature, the team successfully synthesized 18 multilayer phases. Of these, 11 are rare-earth compositions reported for the first time. The findings were published in the journal Science on October 29, 2026, under the title "Data-driven discovery and rapid, direct synthesis of MXenes" (DOI: 10.1126/science.aeh5957).

Technical Details

The discovery process utilized a multi-step workflow rather than autonomous AI design. The team employed machine learning to survey the Materials Project and the Open Quantum Materials Database (OQMD), filtering for atomic arrangements, crystal symmetry, and interatomic distances characteristic of multilayer MXenes. This narrowed the search to 56 experimentally reported compounds, 33 of which were previously classified as layered halide carbides or similar substances.

To produce the materials, the team used self-propagating high-temperature synthesis (SHS). This method involves a brief ignition that triggers a high-temperature chemical reaction finishing within minutes, avoiding the need for sustained external heating. This contrasts with conventional MXene production, which typically requires etching MAX phase precursors with hydrofluoric acid.

Background & Context

MXenes are transition-metal carbides and nitrides known for high electrical conductivity and large specific surface areas, making them suitable for catalysts, electromagnetic shielding, and battery electrodes. The study specifically targeted multilayer structures with the general formula M2XT2 (where M is a metal, X is carbon or nitrogen, and T is a terminating element). Because these structures were studied in the 1970s—before the rise of graphene and 2D materials research—their potential as MXenes was previously unexplored.

Significance & Outlook

The study demonstrates that revisiting legacy crystallographic data through a modern lens can uncover viable materials without starting from scratch. By integrating AI-driven classification with first-principles calculations and rapid synthesis, the researchers established a full workflow for the efficient discovery of 2D materials.

Source: https://xenospectrum.com/en/ai-rediscovered-multilayer-mxenes-science-2026/

Collected: October 31, 2026 | Automated Research System (AI)

#02 AI Materials Models: New Benchmark for Physics-Aware Accuracy

Published October 27, 2026 BIOENGINEER USA

OVERVIEW

Researchers from Columbia Engineering and the University of Cambridge have developed a benchmark to determine if machine learning models for materials science derive macroscopic properties from correct underlying physics. The study reveals that models can predict thermal conductivity accurately while remaining wrong about the atomic vibrations that cause it, highlighting a critical reliability gap for industrial adoption.

IN DEPTH

Key Findings

A new benchmark published in Nature Communications demonstrates that machine learning (ML) models can produce sensible predictions for material properties for the wrong physical reasons. Specifically, the research found that some models with similar accuracy in predicting formation energies and material stability differed significantly in their ability to predict atomic vibrational properties and macroscopic thermal conductivity.

Technical Details

The benchmark evaluates machine-learning interatomic potentials, which are designed to translate quantum-level atomic interactions into macroscopic thermal and mechanical properties. These models learn from datasets of atomic positions, energies, forces, and stresses to avoid solving the Schrödinger equation, allowing them to run roughly 1000 times faster than traditional computational approximations.

The research team, including Michele Simoncelli and Balázs Póta, tested several models against traditional quantum-mechanical calculations across more than 100 different crystalline materials. The team argues that "physics-aware" models must correctly describe atomistic physics—specifically atomic vibrations—to reliably predict macroscopic properties. They noted that judging models solely on energy predictions can overlook errors in forces, which are essential for predicting how materials respond to temperature or mechanical perturbations.

Background & Context

Connecting the microscopic world of nuclei and electrons to tangible engineering properties is a central challenge in computational materials science. While the Schrödinger equation provides the foundation for these calculations, it is too complex to solve analytically for realistic materials. This complexity has led technology companies and startups to adopt ML models to screen candidate materials and automate synthesis via robotic laboratories.

Significance & Outlook

As ML models move from laboratory curiosities to practical tools for material discovery, the question of trustworthiness becomes a practical necessity. The benchmark exposes a gap where speed does not equate to correctness, warning that models may distort the small effects of quantum vibrations while appearing accurate on a larger scale.

Source: https://bioengineer.org/new-benchmark-tests-whether-ai-materials-models-get-physics-right/

Collected: October 31, 2026 | Automated Research System (AI)

#03 Transition Metal Carbides: Composition-Based Melting Point Prediction

Published October 29, 2026 BIOENGINEER

OVERVIEW

Machine learning analysis reveals that elemental composition alone is sufficient to predict the melting temperatures of transition metal carbides. Adding complex descriptors like formation energies and bond geometry produced no statistically significant gain in accuracy, simplifying the data requirements for predicting ultra-high-temperature ceramic properties.

IN DEPTH

Key Findings

A computational study published in the Journal of Materials Science concludes that composition and elemental property descriptors carry the majority of the reproducible predictive signal for the melting points of transition metal carbides. Through controlled paired ablation experiments, researchers found that adding formation energy and local metal–carbon bond geometry descriptors did not provide a statistically significant improvement in predictive accuracy. The increments in the coefficient of determination (R^2) were minimal: 0.0138 for formation energy, 0.0090 for bond geometry, and 0.0124 for the combination, with bootstrap confidence intervals including zero.

Technical Details

The research team, including Wentao Zhou, Nan Qu, Wei Zhang, Tao Xia, Yong Liu, Jingchuan Zhu, and Mingyi Tan, analyzed 105 materials from 74 chemical formula groups and 1,182 local metal–carbon bonds. To ensure data integrity, the team reconstructed bond lengths from the Materials Project database, achieving a mean absolute error of 2.6×10^{-5} Å and a root mean square error of 2.9×10^{-5} Å.

The study employed a rigorous nested evaluation grouped by chemical formula to prevent data leakage. Results from this protocol showed:

- Outer-fold predictions achieved R^2 of 0.806 for formation energy and 0.880 for melting point.
- A separate ensemble based on a finite rule formula across 16 partitions reached a grouped out-of-fold R^2 of 0.911.
- The melting point prediction error for this ensemble was 95.1 K (mean absolute error) and 132.4 K (root mean square error).

Background & Context

Predicting the melting temperatures of ultra-high-temperature ceramics is critical for materials science, but researchers have long debated whether richer information—such as electronic structure, bond geometry, or thermodynamic quantities—is necessary to improve machine learning models over simple elemental composition.

Significance & Outlook

The results suggest that while multiscale descriptors may provide conditional interpretive context to help researchers understand why specific carbides melt at certain temperatures, they do not independently improve prediction accuracy. This indicates that composition-based models are highly effective for predicting the thermal limits of these materials.

Source: https://bioengineer.org/machine-learning-shows-composition-alone-predicts-melting-in-transition-metal-carbides/

Collected: October 31, 2026 | Automated Research System (AI)

#04 EdenGNN: Equivariant Neural Network for Faster Charge Density Prediction

Published October 23, 2026 BIOENGINEER USA

OVERVIEW

Researchers have developed EdenGNN, a machine learning framework that predicts the charge density of crystalline materials over an order of magnitude faster than existing generalizable models. This efficiency gain enables more practical large-scale electronic structure calculations and high-throughput materials screening.

IN DEPTH

Key Findings

The Equivariant Density Graph Neural Network (EdenGNN) significantly reduces the computational cost associated with predicting charge density in crystalline materials. According to the study published in Nature Machine Intelligence, the model is more than an order of magnitude faster than previous generalizable charge density approaches. To verify its generalization capabilities, the team developed EdenGNN-Uni, trained on the Materials Project database, which accurately predicted band structures for most unseen non-magnetic materials across a wide chemical space.

Technical Details

EdenGNN predicts both charge density and augmentation occupancies, allowing for electronic structure calculations with projector augmented-wave accuracy. The framework achieves its efficiency through two primary design choices:

- A basis-expansion formulation utilizing fully trainable radial basis functions.
- A Δ-learning strategy implemented to capture charge transfer.

When tested against DFT calculations using representative structures from the GNoME database, the predicted band structures showed a close match. Furthermore, the authors report that this charge density approach demonstrates enhanced robustness compared to universal machine learning Hamiltonian methods.

Background & Context

Charge density is a fundamental component of density functional theory (DFT), the standard method for computing the electronic structures of materials. While DFT is widely used, it is computationally expensive. Previous machine learning models designed to predict charge density in a generalizable manner were limited by their own prohibitive computational costs, creating a bottleneck that EdenGNN is specifically designed to overcome.

Significance & Outlook

The authors argue that these results establish machine learning charge density models as a scalable ab initio method. This advancement makes high-throughput screening in materials discovery and large-scale electronic structure calculations more feasible. However, the researchers note that performance remains dependent on training data; currently, the reported universal model applies specifically to non-magnetic structures.

Source: https://bioengineer.org/equivariant-neural-network-predicts-charge-density-over-an-order-of-magnitude-faster/

Collected: October 31, 2026 | Automated Research System (AI)

#05 NVIDIA Tensor Cores: Speedups for ML Interatomic Potentials

Published October 29, 2026 Blockchain News USA

OVERVIEW

NVIDIA's Tensor Core FP32 methods, specifically BF16x9 for Blackwell and TF32x3 for Ampere, accelerate machine learning interatomic potentials without requiring model retraining. In tests using the Point Edge Transformer (PET), these methods achieved up to 2.5x GEMM acceleration while maintaining numerical stability close to FP32 baselines. This allows for faster molecular dynamics simulations without the convergence failures seen in standard one-pass TF32.

IN DEPTH

Key Findings

NVIDIA reports that its Tensor Core FP32 methods significantly accelerate the Point Edge Transformer (PET), a transformer-based machine learning interatomic potential developed by EPFL's Lab COSMO. On Blackwell GPUs using the BF16x9 path, large PET general matrix multiplications (GEMMs) were accelerated by 2.2x to 2.5x for a 17,280-atom workload. End-to-end molecular-dynamics step speedups ranged from 1.2x to 1.4x for a 192-atom water box and 1.4x to 1.7x for the 17,280-atom system. On Ampere A100 GPUs using TF32x3, large GEMMs saw speedups of 1.7x to 1.9x, with molecular-dynamics step speedups of 1.5x for a 26-million-parameter checkpoint on the larger system and 1.4x for a 255-million-parameter checkpoint on the 192-atom system.

Technical Details

- **BF16x9 (Blackwell):** Decomposes each FP32 operand into three BF16 components and recombines nine products in FP32. This can be enabled via the CUBLAS_EMULATE_SINGLE_PRECISION=1 environment variable.
- **TF32x3 (Ampere):** Splits each FP32 operand into a leading TF32 component and a residual component, evaluating three cross-products and omitting the residual-by-residual term. This implementation uses CUTLASS's OpMultiplyAddFastF32 operator.

Numerical Stability and Validation

Atomistic simulations are sensitive to numerical drift. In a 30 ps NVE simulation of a 192-atom water box, TF32x3 showed bounded energy fluctuations comparable to FP32, whereas one-pass TF32 exhibited more drift. In a structural-relaxation test on a 539-atom Cantor-alloy slab using the FIRE2 optimizer, FP32 converged in 218 steps and TF32x3 in 218 to 222 steps. Conversely, one-pass TF32 failed to converge within 1,000 steps.

Significance and Outlook

The degree of end-to-end gain depends on system size; larger systems expose more GEMM work, while smaller systems are limited by launch overhead and graph operations. NVIDIA notes that these methods allow for acceleration without changing model architecture or retraining, though they should be validated against specific accuracy and stability requirements for each workload.

Source: https://blockchain.news/news/nvidia-blog-tensor-core-speedups-ml-interatomic-potentials

Collected: October 31, 2026 | Automated Research System (AI)

#06 Epoxy Degradation Analysis: PCA-Based Spectral Fingerprinting

Published October 28, 2026 BIOENGINEER

OVERVIEW

Researchers used principal component analysis (PCA) to link the spectral fingerprints of degrading amine-cured epoxy adhesives to the loss of mechanical strength. The study reveals that degradation behavior varies significantly depending on the curing agent used, with some systems showing non-monotonic strength changes.

IN DEPTH

Key Findings

The study demonstrates that unsupervised machine learning via PCA can systematically organize complex spectral data to track the degradation of filler-containing epoxy formulations. A critical finding is that the curing agent dictates the degradation trajectory: while TETA and AMAM systems showed clearly distinguishable factors at 63 °C, the MDA system's water-related and oxidation-related variations were strongly superimposed at the same temperature, only separating at 87 °C.

Technical Details

The researchers tested three curing agents—triethylenetetramine (TETA), p,p'-methylenedianiline (MDA), and POLYMIDE L-4051 (AMAM)—mixed into a base resin from Cemedine Co., Ltd. containing silica, calcium carbonate, and calcium oxide. Specimens were cast as 0.3-millimeter-thick sheets and subjected to accelerated aging: water immersion at 63 °C and 87 °C, and humid air at 120 °C with 100 percent relative humidity.

Analysis involved measuring bulk tensile strength and recording infrared spectra using a handheld FTIR spectrometer. PCA captured approximately 90 percent of the variance across the first three principal components, which mapped to three themes:

- Moisture-related signals (OH stretching region at 3000 to 3600 per centimeter).
- Oxidation-associated processes (carbonyl, CO, and CN vibrations near 1700 and 1100 per centimeter).
- Filler phase signatures (calcium carbonate bands near 700, 900, and 1400 per centimeter).

Background & Context

The AMAM system exhibited a unique non-monotonic strength profile, where tensile strength dropped, partially recovered, and then fell again. The initial drop is attributed to plasticization from water ingress. The subsequent recovery tracked the second principal component, suggesting that interactions between water and amidoamine groups may induce network relaxation. The final strength loss at 87 °C coincided with changes in the calcium carbonate-associated factor.

Significance & Outlook

The sensitivity of the PCA method allowed the detection of anomalies that would otherwise be missed, such as outliers in the AMAM system at 168 hours. By linking raw spectral variations to mechanical failure without model-dependent assumptions, this approach provides a framework for understanding the complex chemistry of adhesive aging in industrial applications.

Source: https://bioengineer.org/research-report-unsupervised-learning-based-analysis-of-degradation-behavior-in-filler-containing-amine-cured-epoxy-formulations/

Collected: October 31, 2026 | Automated Research System (AI)

#07 MARVEL-AI: A*STAR's AI-Driven Materials Discovery Framework

Published October 30, 2026 A*STAR Singapore

OVERVIEW

A*STAR has launched MARVEL-AI to accelerate the transition of advanced materials from discovery to industrial deployment using an AI-enabled closed-loop system. The initiative demonstrates significant throughput gains, such as a 25-fold increase in polymer sample production and 400-fold increase in alloy screening. This framework aims to bridge the gap between laboratory research and real-world application across semiconductors, energy, and chemicals.

IN DEPTH

Key Findings

MARVEL-AI integrates AI-enabled discovery, automated experimentation, and industrial translation to reduce the time and resources required to bring new materials to market. The platform focuses on four primary material classes: metal alloys, ceramics, composites, and polymers and organics. By linking prediction and validation in an iterative loop, the initiative has already achieved substantial performance milestones in throughput and prototyping.

Technical Details

The system employs a closed-loop architecture where AI models identify candidates and guide automated synthesis and testing, with the resulting data feeding back into the AI to refine future exploration. Specific technical achievements include:

- **Polymers:** Automation increased daily output from three sequential samples in an 8-hour window to 75 samples produced in 24 hours via parallel processing, a 25-fold increase.
- **Metal Alloys:** In-house synthesis and characterization allow for the screening of up to 470 unique alloy samples per day, exceeding conventional methods by more than 400 times.
- **Composites:** The framework enabled the development of a hydrogen-carrying capsule prototype that is four times more durable and 50% lighter than previous versions.

Background & Context

Led by the Agency for Science, Technology and Research (A*STAR) and anchored by the A*STAR Institute of Materials Research and Engineering (IMRE), the initiative is supported by METI, EnterpriseSG, EDB, and NRF under the Research, Innovation and Enterprise 2030 (RIE2030) framework. The platform collaborates with Institutes of Higher Learning (IHLs) and industry partners, including Google Cloud for secure research environments and CuspAI for generative AI integration.

Significance & Outlook

MARVEL-AI seeks to move beyond discovery by ensuring materials are validated under real-world conditions and manufactured at scale. Strategic collaborations target high-performance alloys for aerospace and defense, sustainable materials with reduced reliance on critical raw materials, and advanced membrane materials. The goal is to foster high-value startups and attract multinationals to Singapore's materials development ecosystem.

Source: https://www.a-star.edu.sg/news/press-releases/marvel-ai-accelerating-advanced-materials-innovation-from-discovery-to-deployment

Collected: October 31, 2026 | Automated Research System (AI)

#08 CLaC Model: MIT's AI for Crystal Structure and Text Retrieval

Published October 25, 2026 Intelligent Living USA

OVERVIEW

MIT researchers developed CLaC, a vision-language model that aligns 3D crystal structures with natural language descriptions using contrastive learning. By utilizing synthetic narratives to overcome data bias, the model achieved a 95.4 percent accuracy rate in matching structures to compositions in zero-shot tests. This method provides a scalable way to retrieve promising materials for applications such as solid-state battery electrolytes.

IN DEPTH

Key Findings

The CLaC (Contrastive Language-Crystal) model demonstrates a high capacity for retrieving materials based on geometric and semantic properties. In zero-shot tests, the CLaC-Synthetic version matched crystal structures to the correct composition description with 95.4 percent accuracy, significantly outperforming a text-only Gemma baseline (17.8 percent) and an untrained version (13.5 percent). When queried for solid-state battery electrolyte candidates, the model successfully identified materials previously flagged as promising in separate studies.

Technical Details

CLaC is a vision-language model employing a dual-branch architecture: a crystal graph encoder for 3D atomic arrangements and a text encoder for natural language. The system uses contrastive learning to map these two inputs into a shared mathematical space, allowing researchers to retrieve crystals by typing property descriptions rather than specific formulas. The model was trained using approximately 126,000 materials from the Materials Project. Text data was sourced from paragraphs in four million academic papers and synthetic narratives generated by a language model based on computed properties.

Background and Context

Traditional AI for crystals often treats structures as 3D graphs, while large language models rely on text. This creates a gap because many crystal structures exist in computational databases without accompanying prose. Furthermore, academic literature is biased toward a small set of popular materials. CLaC addresses this by using synthetic narratives, which covered about 300 times more diverse crystals than the literature sample. Testing showed that CLaC-Synthetic achieved structured alignment between geometry and text, whereas the version trained only on academic papers (CLaC-Paper) did not.

Significance and Outlook

While the research did not produce a new commercial battery cell, it establishes a method to bypass data scarcity and popularity bias in materials science. By connecting a material's atomic structure to its properties and processing, the model allows for the discovery of materials based on meaning and geometry rather than fixed labels.

Source: https://www.intelligentliving.co/mit-taught-ai-read-crystals-find-better/

Collected: October 31, 2026 | Automated Research System (AI)

#09 Graph Neural Networks: Accelerating Electrocatalyst Design for 2026

Published October 30, 2026 BIOENGINEER.ORG

OVERVIEW

Graph neural networks (GNNs) are replacing traditional trial-and-error and DFT screening to accelerate the discovery of high-performance electrocatalysts for green hydrogen production. By learning material properties directly from atomic structures, GNNs bypass manual feature engineering to identify low-cost alternatives to platinum and iridium.

IN DEPTH

Key Findings

GNNs have successfully identified numerous high-performance materials for the hydrogen evolution reaction (HER) and oxygen evolution reaction (OER). In HER research, the DimeNet-LSCG model identified 12 platinum alloys with activity comparable to Pt(111), including Cu3Pt and FeCuPt2. Other frameworks identified 43 high-performance alloys from 2,973 candidates and four inexpensive candidates (Fe2Sb4, Cu6Sb2, Cu6Sn2, Ni2Sb2) using Equiformer V2. For OER, researchers identified 14 novel single-atom catalysts with theoretical overpotentials lower than commercial IrO2 and identified 122 bulk-stable and 68 nanoscale-stable potential catalysts from 4,119 oxide materials using GemNet-OC.

Technical Details

GNNs represent molecules or crystals as graphs, treating atoms as vertices and chemical bonds or spatial proximity as edges. Through "message passing," atoms aggregate information from neighbors to learn topological and geometric features automatically. The review categorizes GNN architectures into three paradigms:

- **Invariant models:** (e.g., DTNN, SchNet, DimeNet, GemNet) Convert 3D coordinates into scalars for predicting properties like band gap.
- **Equivariant models:** (e.g., EGNN, PaiNN, NequIP, MACE) Handle both scalar and vector features, essential for molecular dynamics and forces.
- **Transformer-based models:** (e.g., Graphormer, TorchMD-Net, SE(3)-Transformer, Equiformer) Combine global context modeling with geometric symmetry.

Background & Context

Green hydrogen production relies on catalysts for HER at the cathode and OER at the anode. Traditional machine learning relied on manual feature engineering—converting structures into numerical descriptors like electronegativity or d-band center—which is time-consuming and requires deep domain expertise. While density functional theory (DFT) provides high-throughput screening, it remains too costly to exhaustively explore all possible material combinations.

Significance & Outlook

By preserving topological integrity and spatial fidelity, GNNs eliminate the bottleneck of manual descriptor design. This shift toward a GNN-driven paradigm, supported by large-scale public databases such as the Open Catalyst Project (OC20), provides a methodological framework to drastically reduce the time and cost of discovering sustainable, high-performance electrocatalytic materials.

Source: https://bioengineer.org/graph-neural-networks-reshape-electrocatalyst-design-for-green-hydrogen/

Collected: October 31, 2026 | Automated Research System (AI)

#10 Open Polymer Datasets: Reuse Quality Audit across 8 Repositories (2020–2025)

Published October 25, 2026 BIOENGINEER.ORG

OVERVIEW

A cross-repository audit of 53,038 open polymer datasets reveals that the choice of repository is the primary determinant of data reusability, outweighing science or publication year. While deposit volumes grew 143% between 2020 and 2025, average documentation quality remained essentially flat, creating a widening gap between data volume and reuse value.

IN DEPTH

Key Findings

The audit concludes that repository identity explains 37.9% of the variance in reuse scores, an effect an order of magnitude greater than the variance caused by sub-domain, publication year, or sustainability framing. A four-tier quality structure emerged: Dryad, Zenodo, and Figshare occupy the high tier (median scores $\ge$ 5); 4TU.ResearchData, Fordatis, and Mendeley Data are in the middle; Harvard Dataverse follows with a median of three; and OSF collapses to a median of one. The most significant contrast exists between OSF and Dryad, where a randomly selected Dryad record outranks an OSF record in 98.2% of pairings.

Technical Details

The study harvested 53,038 records from eight generalist repositories using 74 polymer- and sustainability-related keywords. Each record was evaluated against 11 binary FAIR-aligned criteria. Results show a structural discontinuity in metadata completeness:

- **High Completion (Bibliographic):** DOI (93.2%), rich abstract (87.4%), keywords (86.1%), and licenses (82.9%).
- **Low Completion (Reuse Value):** Measurement-method description (41.3%), synthesis-method description (32.2%), listed variables (21.5%), open-format files (10.3%), identifiable units (8.3%), READMEs (4.3%), and code links (1.2%).

Background and Context

Polymer informatics and machine-learning models increasingly rely on aggregated open datasets. However, polymer researchers lack domain-specific deposition sites; platforms like Polymer Genome, PI1M, and OpenPoly are read-only, while NOMAD and the Materials Project contain less than 0.1% polymer-relevant content. Consequently, researchers use generalist repositories, but the quality of these deposits has not kept pace with volume. Annual deposits rose from 5,905 in 2020 to 14,374 in 2025, yet the mean reuse score only increased from 3.84 to 4.14.

Significance and Outlook

The open polymer-data ecosystem is widening without deepening. The nineteen-fold gap between volume growth and quality growth indicates that the signal-to-noise ratio is not improving. For downstream researchers, the ability to find, understand, and re-execute datasets remains bounded by the specific repository chosen by the original depositor.

Source: https://bioengineer.org/reuse-quality-of-open-polymer-datasets-a-reproducible-cross-repository-audit-2020-2025/

Collected: October 31, 2026 | Automated Research System (AI)

#11 OrgAIcat: ML model for asymmetric organocatalysis selectivity

Published October 27, 2026 BIOENGINEER.ORG China

OVERVIEW

Researchers have developed OrgAIcat, a machine learning framework that predicts enantioselectivity in asymmetric organocatalysis using a manually compiled dataset of 22,554 data points. The model achieves R^2 values of 0.716 for Aldol reactions and 0.774 for Michael additions, providing a fast route to optimizing chiral drug and functional material synthesis.

IN DEPTH

Key Findings

The OrgAIcat framework demonstrates high predictive accuracy for enantioselectivity in two major reaction classes. Using an ExtraTrees regression model, the team achieved mean absolute errors in free energy difference (ΔΔG) of 0.282 kcal/mol for Aldol reactions and 0.265 kcal/mol for Michael additions. These figures correspond to mean absolute errors in enantiomeric excess (ee) of approximately 7.30% and 6.78%, respectively. Additionally, a four-category classification model for selectivity intervals reached AUC values of 0.908 and 0.922.

Technical Details

The model is powered by iSynth, a dataset extracted from 844 papers published between 2000 and 2020. After rigorous cleaning, the final set included 13,360 Aldol reactions and 9,194 Michael addition reactions. To represent these reactions, the team developed the R-SPOC descriptor, which integrates three data streams:

- Catalyst structure and physicochemical characteristics.
- Structural changes (the difference between product and reactant descriptors).
- Reaction conditions, including temperature, solvent, and additives.

R-SPOC eliminates the need for quantum chemical calculations or explicit transition state searches, reducing computational costs.

Background and Context

The researchers identified a significant "selection bias" in existing chemistry literature. While the potential reaction space for Aldol and Michael additions is estimated at 1016 and 1017 combinations respectively, published literature only records approximately 104, heavily skewed toward high-performing reactions. To ensure the model learned structure-selectivity relationships rather than mere popularity, the team performed low-similarity tests and downsampled high-frequency components.

Significance and Outlook

OrgAIcat's utility was validated through external tests on papers published between 2021 and 2024 and laboratory experiments. In a test involving 24 newly synthesized chiral primary amine catalysts, the model tracked selectivity differences with a Pearson correlation coefficient of 0.736. This framework transforms imperfect historical literature into a predictive tool for the production of fine chemicals and chiral drugs.

Source: https://bioengineer.org/machine-learning-model-turns-decades-of-literature-data-into-a-fast-route-to-highly-selective-organocatalysis/

Collected: October 31, 2026 | Automated Research System (AI)

#12 qNEP Potential: Sodium Battery SEI Formation and Simulation Specs

Published October 25, 2026 BIOENGINEER.ORG

OVERVIEW

Researchers have developed a charge-aware neuroevolution potential (qNEP) that extends simulation timescales for sodium battery interfaces to 100 nanoseconds. This machine-learning approach overcomes previous bottlenecks in observing the formation of the solid electrolyte interphase (SEI), enabling the study of nanosecond-scale densification and phase separation.

IN DEPTH

Key Findings

The development of the qNEP (charge-aware neuroevolution potential) allows for the simulation of sodium metal electrode-organic electrolyte interfaces over timescales significantly longer than previous methods. In a system of roughly 1,000 atoms, the potential remained stable for 100 nanoseconds, representing a two to three order of magnitude improvement over prior lithium and sodium simulations. For larger systems of approximately 27,000 atoms, simulations lasting tens of nanoseconds were achieved, with the longest single-GPU run taking eight days.

Technical Details

The qNEP strategy utilizes three primary pillars:

- High-temperature ab initio molecular dynamics sampling to broaden descriptor space coverage and enhance ergodicity.
- Injection of short ab initio simulations at critical nodes during active learning to prevent unphysical energy pathways at bifurcation points.
- A training set incorporating inorganic products (sodium fluoride, sodium oxide) and organic decomposition intermediates.

The model was trained using the GPUMD package on four RTX 4090 GPUs over approximately two days. It contains 22,633 trainable parameters for six elements, based on a training set of 5,035 configurations totaling 683,526 atoms. Accuracy metrics show root mean square errors of 10.50 meV per atom for energy and 251.89 meV per angstrom for forces on the training set. This force error is roughly 38 times smaller than that of the ReaxFF force field (9,680 meV per angstrom). Speed-wise, qNEP is one order of magnitude faster than ReaxFF and six orders of magnitude faster than ab initio molecular dynamics.

Background & Context

Sodium batteries offer a cheap, abundant alternative to lithium-ion systems, but their performance is limited by the solid electrolyte interphase (SEI)—a passivation film formed during initial charging. Direct observation of SEI formation is difficult because conventional techniques like X-ray photoelectron spectroscopy provide only snapshots. Previous simulations were either too short (ab initio) or lacked the transferability and accuracy to capture single-electron reduction pathways (ReaxFF).

Significance & Outlook

By incorporating trainable atomic charges, qNEP can track charge transfer and distinguish chemical states, a feature missing in standard machine learning potentials. This allows researchers to simulate and differentiate the decomposition pathways of various electrolytes, including carbonates (ethylene carbonate, dimethyl carbonate, propylene carbonate) and ethers (dimethoxyethane, diglyme).

Source: https://bioengineer.org/machine-learning-simulations-reveal-how-sodium-batteries-form-their-protective-interface/

Collected: October 31, 2026 | Automated Research System (AI)

#13 Lithium Metal Batteries: AI-Driven Discovery of 368 Non-Fluorinated Diluents

Published October 29, 2026 AI타임스 South Korea

OVERVIEW

Researchers at Seoul National University developed an AI algorithm that screened 1.32 million molecules to identify 368 promising non-fluorinated electrolyte diluents for lithium metal batteries. Three of these candidates were experimentally verified in pouch cells, demonstrating stable charge-discharge performance. This approach significantly reduces the time and cost of material discovery by replacing trial-and-error synthesis with physics-based machine learning.

IN DEPTH

Key Findings

A research team led by Professor Jong-Woo Lim at Seoul National University has identified 368 promising non-fluorinated electrolyte diluent candidates from a database of approximately 1.32 million molecules. By applying a machine learning algorithm to the PubChem database, the team narrowed the search space to approximately 0.028% of the original pool. To validate these results, the researchers applied three of the selected materials to lithium metal single-layer pouch cells, confirming stable operation during repeated charging and discharging cycles.

Technical Details

The AI algorithm utilizes two specific physical and chemical descriptors to predict the suitability of a molecule as a diluent:

- **BF_3 Affinity:** Measures the electron-pair donating characteristics to evaluate the molecule's tendency to bind directly with lithium ions.
- **Molecular Polarity Index (MPI):** Analyzes charge distribution and electrostatic interactions to determine how well the diluent interacts with other electrolyte components.

Unlike standard AI models that rely solely on statistical patterns, this system uses these chemical criteria to ensure the diluent does not bind strongly to lithium ions while maintaining appropriate interactions with the rest of the electrolyte.

Background and Context

Lithium metal batteries offer higher energy density than lithium-ion batteries but suffer from high chemical reactivity, leading to side reactions that degrade stability and lifespan. Localized High-Concentration Electrolytes (LHCE) address this by adding diluents to high-concentration electrolytes to lower viscosity without disrupting the lithium-ion coordination structure. While fluorinated compounds are commonly used, there is a demand for non-fluorinated alternatives due to high synthesis costs and environmental persistence concerns.

Significance and Outlook

This research demonstrates that physics-grounded AI can drastically compress the experimental timeline for material discovery. While only three candidates have been experimentally verified so far, the study expands the library of viable non-fluorinated alternatives. Future commercialization will require further evaluation of long-term cycle life, stability under various operating conditions, and mass-production feasibility.

]]

Source: https://www.aitimes.kr/news/articleView.html?idxno=42475

Collected: October 31, 2026 | Automated Research System (AI)