INTERNATIONAL JOURNAL OF ADVANCED RESEARCH
International Peer-Reviewed Research Journal
ISSN (Online): 3106-7980 | ISSN (Print): 3106-7972
Volume 3 • Issue 7 • November 2026

ARTIFICIAL INTELLIGENCE AND DENSITY FUNCTIONAL THEORY DRIVEN DESIGN OF ADVANCED FUNCTIONAL MATERIALS FOR SUSTAINABLE ENERGY CONVERSION AND STORAGE

Kalsoom Panhwar[1], Imran Khan Swati[2], Jalal Usmani[3], Shayan Shakir[4], Syed Mazahir Hussain[*5], Asad Ali Shaikh[6]

[1]Department of Information Technology, University of Sindh,
[2]Department Of Chemical Engineering, University of Engineering and Technology Peshawar Pakistan
[3]Lubricant and Chemical Department, Pakistan State Oil, Karachi, Pakistan
[4]Department of Mechanical Engineering & Technology, Hamdard University, Karachi, pakistan
*[*5]Department of Material Science and Engineering, Institute of Space Technology, Islamabad*
[6]Department of Mathematics, Sir Syed University of Engineering and Technology,Karachi Pakistan

[1]kulsoom.panhwar@usindh.edu.pk, [2]imranswati@uetpeshawar.edu.pk, [3]jalal_usmani_909@yahoo.com, shayan.shakirr@gmail.com, [*5]syedmazahir2533@gmail.com, [6]asadalishaikh711@gmail.com,

ORCID ID: [1]0009-0004-9396-899X, [2]0009-0008-6314-1392, [4]0009-0005-3846-4550, [5]0009-0009-3642-7798, [6]0009-0008-3831-6015

DOI: https://doi.org/10.5281/zenodo.24000950

Keywords
density functional theory; machine learning; graph neural networks; high-throughput screening; inverse design; batteries; solid electrolytes; perovskite solar cells; electrocatalysis; energy materials

Article History
Received: 06 August 2026
Accepted: 20 October 2026
Published: 05 November 2026

Copyright @Author
Corresponding Author: *
Syed Mazahir Hussain
syedmazahir2533@gmail

Abstract

The global transition to a low-carbon energy system depends on materials that can harvest, convert and store energy with high efficiency, long lifetime and low cost. Conventional discovery based on empirical trial and error is too slow to meet this demand, because the chemical and structural search space of inorganic, hybrid and molecular materials is astronomically large. Density functional theory (DFT) has become the standard quantum-mechanical tool for predicting the stability, electronic structure, ionic transport and surface reactivity of candidate materials, yet its computational cost limits the number of systems that can be examined. Artificial intelligence (AI), particularly machine learning (ML) with graph neural networks, machine-learned interatomic potentials and generative models, can learn from DFT data and extend its reach by several orders of magnitude. This article reviews how the combination of DFT and AI is reshaping the design of functional materials for batteries, solid-state electrolytes, perovskite photovoltaics, photocatalysts and electrocatalysts. We summarize the theoretical foundations and known limitations of DFT, the main families of ML models and databases, and an integrated closed-loop workflow in which DFT, ML, active learning and experiment reinforce one another. Representative case studies are analyzed, and open challenges, namely functional accuracy, data quality, out-of-distribution generalization, synthesizability, interpretability and experimental validation, are discussed together with future directions.

1. Introduction

Meeting global climate targets requires a rapid expansion of renewable electricity generation together with large-scale, affordable energy storage and clean chemical fuels. Each of these technologies is limited, at its core, by the properties of a material: the light absorber in a solar cell, the intercalation host in a battery

electrode, the ionic conductor in a solid electrolyte, the electrode of a supercapacitor, or the catalyst that drives water splitting and carbon dioxide reduction [1, 2, 3, 4]. Progress in lithium-ion batteries illustrates the time scale of conventional materials development: decades passed between the identification of layered oxide cathodes and their widespread commercial deployment, and further improvements in energy density, safety and cost still depend on finding new compositions and structures [1, 3]. A systematic, accelerated route to new materials is therefore a scientific and an economic priority.

The foundation of such a route was laid by density functional theory. The Hohenberg–Kohn theorems and the Kohn–Sham equations showed that the ground-state properties of a many-electron system can be obtained from the electron density rather than the many-body wavefunction [5, 6]. Together with efficient plane-wave and pseudopotential implementations [7, 8, 9], this made it possible to compute formation energies, voltages, band gaps, migration barriers and adsorption energies from the atomic structure alone. As early as the late 1990s, first-principles calculations were used to predict lithium intercalation voltages and to guide the search for new cathode materials [10, 11]. The subsequent "materials genome" approach scaled this idea by automating thousands of calculations and sharing the results in open databases such as the Materials Project, AFLOW, OQMD and JARVIS [12, 13, 14, 15]. High-throughput DFT nevertheless remains expensive. A single hybrid-functional calculation on a supercell with a few hundred atoms can require thousands of core-hours, and the number of plausible multi-element compounds, surfaces, interfaces and defect configurations far exceeds what can be computed directly. Machine learning offers a complementary capability. Models trained on DFT data can predict properties in milliseconds, learn interatomic forces with near-DFT accuracy, rank candidates for experimental follow-up, and, more recently, generate new crystal structures conditioned on target properties [16, 17, 18, 19]. The two methods are mutually dependent: ML needs the accurate, physically consistent labels that DFT provides, while DFT gains reach, speed and guidance on where to compute next.

The purpose of this article is to provide a coherent account of how DFT and AI are combined in the design of functional materials for sustainable energy. Section 2 summarizes the DFT methodology and the energy-relevant quantities it provides, including its main limitations. Section 3 reviews the AI toolbox: databases, descriptors, graph neural networks, machine-learned interatomic potentials, generative models and active learning. Section 4 presents an integrated closed-loop workflow. Section 5 analyzes applications in batteries, solid electrolytes, photovoltaics, photocatalysis and electrocatalysis. Section 6 discusses the principal challenges and outlook, and Section 7 concludes.

2. Density Functional Theory for Energy Materials

2.1 Theoretical background

In Kohn–Sham DFT the interacting electron system is mapped onto a fictitious non-interacting system that reproduces the same ground-state density $n(\mathbf{r})$. The single-particle orbitals satisfy [6]

$$[-(\hbar^2/2m)\nabla^2 + v_{eff}(\mathbf{r})]\,\psi_{\langle i\rangle i\langle/i\rangle}(\mathbf{r}) = \varepsilon_{\langle i\rangle i\langle/i\rangle}\,\psi_{\langle i\rangle i\langle/i\rangle}(\mathbf{r}) \quad (1)$$

with the effective potential $v_{eff} = v_{ext} + v_H + v_{xc}$, where the external, Hartree and exchange–correlation contributions appear, and the density is $n(\mathbf{r}) = \Sigma_{\langle i\rangle i\langle/i\rangle}\,|\psi_{\langle i\rangle i\langle/i\rangle}(\mathbf{r})|^2$. The exchange–correlation functional $E_{xc}[n]$ is the only quantity that must be approximated, and the quality of every DFT prediction rests on this approximation. The generalized gradient approximation of Perdew, Burke and Ernzerhof (PBE) is the most widely used choice for solids because it balances accuracy and cost [20]. Meta-GGA functionals such as SCAN improve the description of diverse bonding environments [21], and screened hybrid functionals such as HSE06 mix a fraction of exact exchange to correct band gaps and localization errors [22]. For adsorption on metal surfaces, revised functionals such as RPBE are often preferred [23].

2.2 Known limitations and common corrections

Because energy materials involve transition-

metal oxides, weakly bound layers, charged defects and solid–liquid interfaces, several systematic limitations of semilocal DFT must be handled with care:

- **Band-gap underestimation.** Semilocal functionals typically underestimate band gaps substantially. Hybrid functionals or many-body methods are required for quantitative optical and photocatalytic predictions [22].
- **Self-interaction error in localized d and f states.** The DFT+U approach adds an on-site Hubbard term to correct the delocalization of transition-metal electrons and is essential for reliable redox potentials of cathode materials [24, 25]. Mixing GGA and GGA+U calculations in a single thermodynamic framework requires explicit energy corrections [26].
- **Dispersion interactions.** Van der Waals forces, important for layered materials, intercalation and molecule–surface binding, are absent from standard functionals and are commonly added through pairwise corrections such as DFT-D3 [27].
- **Solvation and electrochemical environment.** Reactions at electrified interfaces require implicit or explicit solvent models and a treatment of the electrode potential; simplified thermodynamic models, such as the computational hydrogen electrode, are widely used [28].
- **Length and time scales.** Ab initio molecular dynamics is restricted to a few hundred atoms and picoseconds to nanoseconds, which is insufficient for slow ion hopping, grain boundaries or long-term degradation.

2.3 Energy-relevant quantities obtained from DFT

A principal strength of DFT is that physically meaningful figures of merit can be computed from total energies. The thermodynamic stability of a compound is measured by its energy above the convex hull, E_{hull}, i.e., the energy released by decomposition into the most stable competing phases; compounds with small E_{hull} are candidates for synthesis, and the thermodynamic scale of metastability of known inorganic compounds has been mapped systematically [29]. For an intercalation electrode, the average voltage between two lithium contents x_1 and x_2 follows from the total energies of the lithiated and delithiated hosts and of metallic lithium [10]:

$$V = -[E(Li_{<i>x</i>_2}Host) - E(Li_{<i>x</i>_1}Host) - (x_2 - x_1)E(Li)] / [(x_2 - x_1)e] \quad (2)$$

Ionic conductivity is controlled by the migration barrier E_a, which is obtained with the nudged elastic band method [30] or from ab initio molecular dynamics, and enters the Arrhenius law $\sigma = (\sigma_0/T)\exp(-E_a/k_BT)$. For electrocatalysis, the free energy of an adsorbed intermediate is evaluated as [28]

$$\Delta G = \Delta E_{ads} + \Delta E_{ZPE} - T\Delta S \quad (3)$$

where E_{ZPE} is the zero-point energy and S the entropy. For photovoltaic and photocatalytic absorbers, the band gap, band-edge positions, effective masses and defect formation energies are the key quantities. Table 1 summarizes the main DFT-derived descriptors used for each class of energy material.

Table 1. Representative DFT-derived descriptors for energy-conversion and storage materials.

Application	Key descriptors	Typical DFT method
Li/Na-ion cathodes	Average voltage, volume change, E_{hull}, migration barrier	GGA+U total energies, NEB [10, 31]
Solid electrolytes	Li-ion migration barrier, electrochemical window, interface stability	AIMD, NEB, grand-potential phase diagrams [32, 33]
Photovoltaic absorbers	Band gap, effective masses, defect levels, absorption	HSE06 (with spin–orbit coupling), defect supercells [22, 34]
Photocatalysts	Band-edge alignment, gap, surface stability	Hybrid functionals, slab calculations [35]

Application	Key descriptors	Typical DFT method
Electrocatalysts	Adsorption free energies of intermediates, overpotential	RPBE/PBE slabs, computational hydrogen electrode [23, 28]

3. Artificial Intelligence Methodologies in Materials Design

3.1 Databases and high-throughput infrastructure

The availability of large, curated datasets is the precondition for data-driven materials science. The Materials Project, AFLOW, OQMD and JARVIS-DFT provide DFT-computed structures, energies, electronic structures and elastic or optical properties for well over a hundred thousand inorganic compounds [12, 13, 14, 15]. Software ecosystems such as pymatgen standardize structure handling, input generation and analysis [36]. For heterogeneous catalysis, the Open Catalyst datasets contain millions of DFT relaxations of adsorbates on surfaces, with OC20 focused on metals and intermetallics and OC22 on oxide electrocatalysts [37, 38]. Text-mining of the literature offers an additional, non-DFT data source: unsupervised word embeddings trained on millions of abstracts were shown to capture latent materials knowledge and to suggest candidate thermoelectrics before their reporting [39].

3.2 Descriptors and classical machine learning

Early ML studies represented a material by hand-crafted descriptors, such as elemental properties, stoichiometric statistics and structural fingerprints, and fitted regression models, including kernel ridge regression, random forests and Gaussian processes, to DFT labels [40, 41, 42]. The choice of descriptor strongly determines model quality and transferability, motivating methods that search for compact, physically interpretable descriptors; the SISSO approach, for example, applies compressed sensing to a large space of analytic expressions and returns low-dimensional formulas for target properties [43, 44]. These models are data-efficient and interpretable, which makes them attractive when only hundreds to thousands of calculations are available.

3.3 Graph neural networks

Graph neural networks (GNNs) represent a crystal as a graph whose nodes are atoms and whose edges encode interatomic distances, and they learn the representation directly from the structure. The crystal graph convolutional neural network (CGCNN) demonstrated accurate prediction of formation energies, band gaps and elastic moduli, and offered chemical insight by decomposing a property into site contributions [45]. MEGNet generalized the idea with global state variables to treat molecules and crystals in one framework [46], SchNet introduced continuous-filter convolutions [47], and ALIGNN added bond-angle information through a line graph [48]. Systematic benchmarks show that the accuracy ranking of architectures depends on the target property and dataset size, and that data quality is often as important as model design [49]. A recurring caveat is that models predicting formation energy with small average errors can still be unreliable for stability, because stability depends on small energy differences between competing phases [50]; structure-based networks that predict stability more directly have been proposed to mitigate this problem [51].

3.4 Machine-learned interatomic potentials and universal models

Machine-learned interatomic potentials (MLIPs) replace the quantum-mechanical energy evaluation by a flexible function fitted to DFT energies and forces. Early examples include high-dimensional neural-network potentials and Gaussian approximation potentials [52, 53], and the field has been reviewed extensively [18, 54]. Equivariant architectures such as NequIP and MACE enforce the rotational symmetry of physical interactions and reach high accuracy from small training sets [55, 56]. A major recent development is the universal potential, trained on very large and chemically diverse datasets so that it can be applied across the periodic table without retraining: M3GNet and CHGNet, trained on Materials Project relaxation trajectories, allow structure relaxation and molecular dynamics at a small fraction of the

DFT cost, and CHGNet additionally incorporates magnetic moments as a proxy for charge state [57, 58]. Such potentials make it feasible to study ion diffusion, phase transformations and interfaces on nanosecond time scales, though their accuracy near transition states and for out-of-distribution chemistries still requires verification against DFT.

3.5 Generative models and inverse design

Forward models map structure to property; inverse design aims at the opposite mapping, from a target property to a structure [59, 60]. Generative models learn the distribution of known stable crystals and sample new ones. The crystal diffusion variational autoencoder (CDVAE) generates periodic structures with physically plausible local geometries [61], and diffusion models such as MatterGen can be conditioned on chemistry, symmetry and target properties, with DFT used to verify that a fraction of the generated materials are stable [19]. At a larger scale, the GNoME project combined GNN-based filtering with iterative DFT verification to report approximately 2.2 million new crystal structures predicted to be stable, of which about 380,000 lie on the updated convex hull [62]. Complementary ab initio structure-search methods, such as random structure searching and evolutionary algorithms, remain important for validating generated structures [63, 64].

3.6 Active learning, Bayesian optimization and autonomous experimentation

When each evaluation is expensive, the sequence of calculations should itself be optimized. Bayesian optimization builds a probabilistic surrogate of the objective and selects the next candidate by balancing exploitation and exploration [65]. Active learning applies the same principle to dataset construction by choosing the points where the model is most uncertain, which improves data efficiency [66]. These methods link naturally with robotic synthesis and characterization: the A-Lab platform reported the autonomous synthesis of 41 of 58 target inorganic compounds in 17 days using computed targets and literature-trained recipe proposals, although some structural assignments in that report have been questioned, which underlines the need for independent validation [67]. Self-driving laboratories are widely viewed as a route to closing the loop between computation and experiment for clean-energy materials [68].

3.7 Interpretability and uncertainty

For scientific use, a prediction must be accompanied by an explanation and an estimate of its reliability. Attribution methods such as SHAP quantify the contribution of each input feature to a prediction [69], and the site-wise decomposition of CGCNN gives a chemical reading of the output [45]. Ensembles, Bayesian neural networks and Gaussian processes provide uncertainty estimates that can drive active learning and flag extrapolation. Table 2 lists the principal model families together with their typical use and main limitation.

Table 2. Main families of AI models used in DFT-driven materials design.

Model family	Typical use	Strengths	Limitations
Descriptor-based ML (RF, kernel, GP)	Property regression and screening [40, 41]	Data-efficient, interpretable	Descriptor dependent, limited transfer
Graph neural networks (CGCNN, MEGNet, ALIGNN)	Crystal-property prediction [45, 46, 48]	Learn features from structure	Need relaxed structures, large data

Model family	Typical use	Strengths	Limitations
Interatomic potentials (NequIP, MACE, M3GNet, CHGNet)	Relaxation, MD, diffusion [55, 57, 58]	Near-DFT accuracy at low cost	Reliability near transition states and for new chemistries
Generative models (CDVAE, MatterGen)	Inverse design of crystals [19, 61]	Propose novel structures	Synthesizability not guaranteed
Bayesian optimization / active learning	Efficient search and data acquisition [65, 66]	Reduces number of DFT calls	Depends on surrogate and acquisition design

4. An Integrated AI–DFT Workflow

The methods described above are most powerful when combined into a closed-loop workflow, outlined in Fig. 1. The process begins with a precise definition of the target: the application, the figures of merit (for example, voltage and capacity for a cathode, or Li-ion barrier and electrochemical window for an electrolyte) and the constraints on elemental abundance, toxicity and cost. Existing data from open databases are then combined with new high-throughput DFT calculations to build a training set whose labels are computed with a consistent level of theory [12, 26]. Surrogate models, GNNs for properties and MLIPs for dynamics, are trained on these labels and used to screen a candidate space that may contain millions of compounds, or to guide a generative model toward promising regions [19, 62].

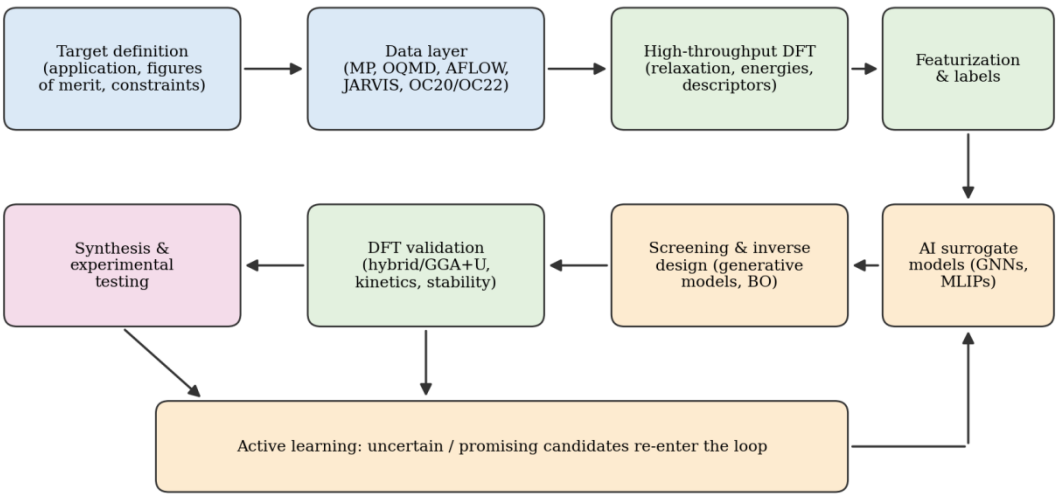


Fig. 1. Integrated closed-loop AI–DFT workflow for materials design. Surrogate models screen and generate candidates; high-fidelity DFT validates the shortlist; experimental feedback and model uncertainty drive active learning.

The shortlist produced by the surrogate stage is then re-examined at higher fidelity. This step typically involves hybrid-functional or GGA+U calculations, explicit evaluation of E_{hull} and decomposition products, calculation of migration barriers or adsorption free energies, and analysis of interface and electrochemical stability. Only candidates that pass these physical filters are proposed for synthesis. The new DFT results, and later the experimental results, are fed back into the training set, and active learning identifies regions of chemical space where the model is uncertain and where additional calculations will be most informative [66]. In this

way, DFT acts as the physical "oracle" that keeps the AI models anchored to quantum-mechanical reality, while the AI models decide where the expensive oracle should be consulted.

Two design principles make such a workflow robust. First, multi-fidelity learning, in which large amounts of inexpensive GGA data are combined with small amounts of hybrid-functional or experimental data, reduces the cost of reaching high accuracy. Second, the data split used to evaluate models should reflect the intended use: random splits tend to overestimate performance, whereas splits by chemical system or by time better indicate how a model behaves when asked to predict genuinely new materials [50].

5. Applications in Sustainable Energy Conversion and Storage

5.1 Electrode materials for lithium-ion and sodium-ion batteries

Battery electrodes were among the first successes of first-principles materials design. Ab initio calculations of lithium intercalation voltages in oxides and chalcogenides reproduced experimental trends and demonstrated that voltage could be predicted before synthesis [10]. Subsequent work showed how computed thermodynamic and kinetic quantities can be used to identify new cathode chemistries [11, 70]. A systematic high-throughput evaluation of lithium transition-metal phosphates computed voltage, stability and safety-related properties for a large set of compositions and prototype structures and singled out several promising families [71]. The same framework was applied to sodium-ion systems, where the calculations showed systematic differences from lithium in voltage, phase stability and diffusion barrier, and thereby guided the choice of host structures for Na intercalation [31]. Comprehensive reviews of computational battery research summarize the accuracy that can be expected for voltages, phase diagrams and diffusion [72].

Machine learning extends this paradigm in three ways. First, GNN property models trained on DFT voltages, formation energies and volume changes enable screening of much larger spaces of compositions and structures than direct DFT permits [45, 73]. Second, universal MLIPs allow relaxation of partially lithiated configurations and molecular dynamics of lithium diffusion in structures that would be too costly for DFT [57, 58]. Third, generative and active-learning strategies can propose new framework structures while minimizing the number of DFT calculations required. Because redox voltages depend sensitively on the Hubbard U parameter and on charge ordering, all ML predictions for transition-metal electrodes should be understood as inheriting the systematic errors of the functional on which they were trained [25, 26].

5.2 Solid-state electrolytes

All-solid-state batteries promise higher safety and, with lithium-metal anodes, higher energy density, but they require solid electrolytes combining high ionic conductivity with a wide electrochemical stability window and compatible interfaces. The discovery of the lithium superionic conductor $Li_{10}GeP_2S_{12}$, with conductivity comparable to that of liquid electrolytes, was a landmark [74]. First-principles calculations subsequently explained its diffusion pathways and its limited electrochemical window [32]. A broader analysis of candidate materials showed that a body-centered-cubic arrangement of anions is associated with low migration barriers, providing a structural design principle for superionic conductors [33].

Data-driven screening has turned such principles into systematic searches. Sendek and co-workers screened more than 12,000 lithium-containing structures with a sequence of filters, including stability, electronic insulation and a ML model trained on known conductors, and identified a small set of promising candidates for further study [75]. Universal MLIPs now make it possible to estimate diffusion coefficients through molecular dynamics for hundreds of candidates at DFT-like fidelity, which was previously impractical [57, 58]. Important open issues include the stability of the electrolyte–electrode interface, grain-boundary resistance and the coupling of mechanical and chemical degradation, which are difficult to represent in small periodic cells and which call for multiscale approaches that combine MLIPs with continuum models.

5.3 Perovskite and other photovoltaic absorbers

The efficiency of a single-junction solar cell is fundamentally limited by the band gap of the absorber, with a maximum of roughly one third near a gap of 1.3–1.4 eV in the detailed-balance model [76]. Halide perovskites entered photovoltaics as sensitizers in 2009 [77] and have since reached power conversion efficiencies above 25% in single-junction cells. DFT calculations helped to explain this performance: the unusual defect physics of methylammonium lead iodide, in which the dominant native defects introduce shallow rather than deep levels, rationalized the long carrier lifetimes of these materials [34]. The concept of defect tolerance has since been generalized into design rules for finding other compounds with similarly benign defect properties [78].

Two practical obstacles motivate computational and AI-driven searches: the toxicity of lead and the chemical instability of hybrid organic–inorganic compounds. Machine learning has been used to screen thousands of candidate lead-free hybrid perovskites for stability and band gap, yielding a short list of compositions that was subsequently examined by DFT [79]. In such studies, DFT with hybrid functionals and spin–orbit coupling is required to obtain trustworthy band gaps, whereas ML models trained on large GGA datasets provide a fast pre-filter [22]. The principal challenge in the perovskite field is that long-term stability under heat, humidity and illumination depends on kinetic degradation pathways and interfaces, properties that are poorly captured by current surrogate models and that require MLIP-based simulation of realistic interfaces.

5.4 Photocatalysts for water splitting

Since the demonstration of photoelectrochemical water splitting on TiO_2 [35], the search for photocatalysts that absorb visible light, have band edges straddling the water redox potentials and remain stable in aqueous solution has been a central problem. DFT provides the band gap and band-edge positions, while the stability criterion requires comparison of the material's decomposition potentials with the water redox levels. This multi-criteria problem is well suited to high-throughput screening: the band gap and edge positions can be predicted by GNNs trained on hybrid-functional data, and the surface stability can be checked with surface phase diagrams computed by DFT or MLIP. Because the semilocal band-gap error is large, the quality of the underlying training labels is critical, and it is advisable to use hybrid or GW data, or to apply a systematic correction, when training ML models for this purpose [22].

5.5 Electrocatalysis: hydrogen evolution, oxygen evolution and CO₂ reduction

Electrocatalysis is the area in which DFT-based descriptor thinking has been most successful. The d-band model related the adsorption strength of molecules on transition-metal surfaces to the electronic structure of the metal [80]. For the oxygen reduction reaction, DFT-computed adsorption free energies of intermediates explained the origin of the overpotential and produced a volcano relationship between activity and the oxygen binding energy [28]. For hydrogen evolution, the hydrogen adsorption free energy ΔG_{H^*} near zero is the descriptor of an ideal catalyst (Fig. 2), and computational screening of surface alloys in this way identified a bismuth–platinum surface alloy that was subsequently confirmed experimentally [81, 82]. For oxygen evolution on oxides, a scaling relation between the adsorption energies of OH* and OOH* sets a minimum theoretical overpotential of roughly 0.4 V for many materials, and this has become a standard benchmark in the search for improved oxygen-evolution catalysts [83, 84].

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH
International Peer-Reviewed Research Journal
ISSN (Online): 3106-7980 | ISSN (Print): 3106-7972
Volume 3 • Issue 7 • November 2026

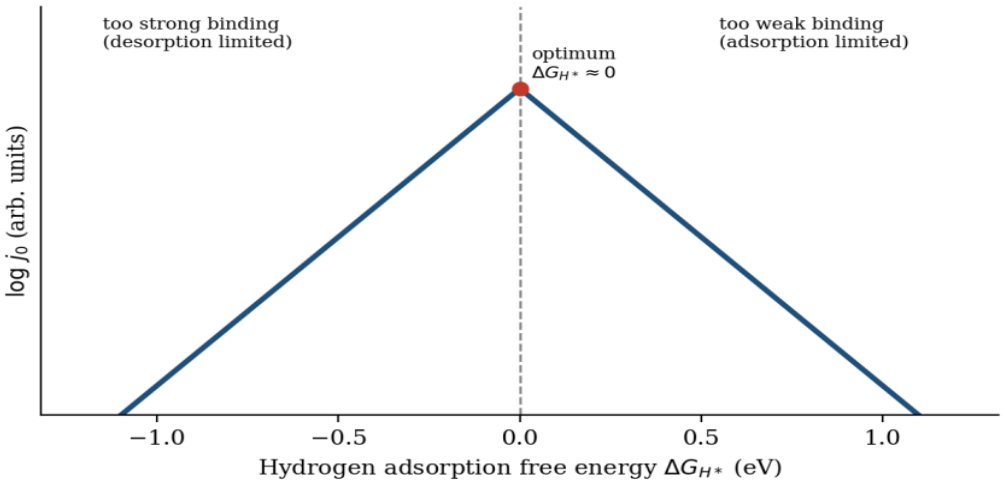


Fig. 2. Schematic volcano relationship for the hydrogen evolution reaction (HER). Activity is maximal when the hydrogen adsorption free energy is close to zero; weaker or stronger binding limits the rate (after the concept of [81, 82]).

For CO_2 reduction on copper, DFT identified the key intermediates and the binding energy of CO as a determinant of the product selectivity [85]. This system also illustrates the added value of ML: Zhong et al. combined DFT-computed descriptors with active machine learning to select among a large set of copper-based alloy candidates, and the experimentally verified best catalyst showed high selectivity toward ethylene [86]. The Open Catalyst datasets have made it possible to train GNNs that predict relaxed adsorption energies directly from the initial structure, thereby replacing thousands of DFT relaxations per candidate surface [37, 38]. Reviews of the combined theory–experiment approach emphasize that predicted descriptor trends should always be validated under realistic potential, pH and coverage conditions, as these significantly alter the active site [87].

Table 3 summarizes the representative studies discussed in this section and the role played by DFT and AI in each.

Table 3. Selected case studies illustrating the roles of DFT and AI in energy-materials design.

Application	Contribution	Main method	Ref.
Li-ion cathodes	Voltage prediction and screening of phosphate cathodes	GGA+U high-throughput	[10, 71]
Solid electrolytes	Screening of >12,000 Li-containing structures	DFT + ML filters	[75]
Superionic conductors	Anion-framework design principle for low barriers	DFT/AIMD analysis	[33]
Perovskite PV	Defect tolerance; ML screening of lead-free hybrids	Hybrid DFT, ML	[34, 79]
HER catalysts	Computational screening of surface alloys	DFT descriptor	[81]
CO_2 reduction	Active-learning discovery of alloy catalysts	DFT + active ML	[86]
Stable crystals	2.2 million predicted structures	GNN + DFT loop	[62]

6. Challenges and Future Perspectives

Accuracy and consis of the underlying DFT data. Machine-learning models cannot be more accurate than the labels they are trained on. Large public datasets are generated mainly with GGA-level functionals and, for oxides, with a single Hubbard U value, which introduces systematic errors for band gaps, redox potentials and adsorption energies [20, 25]. Combining datasets computed with different settings can introduce inconsistencies unless corrections are applied [26]. Transfer learning and multi-fidelity learning, in which models are pretrained on abundant low-level data and then fine-tuned on scarce high-level or experimental data, are promising ways forward, as are meta-GGA and hybrid datasets of increasing size [21].

Generalization and out-of-distribution prediction. The ultimate goal of discovery is to find materials unlike those already known, but ML models are most reliable close to their training distribution. Evidence that models with low average error may still fail at identifying stable compounds highlights the need for realistic benchmarks and for uncertainty quantification [50]. Universal potentials in particular should be validated systematically for properties that depend on the curvature of the potential energy surface and for transition states [57, 58].

Synthesizability and kinetics. Thermodynamic stability is necessary but not sufficient for synthesis. Many computationally predicted compounds may be difficult or impossible to make because of kinetic barriers or competing phases, and the thermodynamic scale of metastability shows that many known materials are themselves metastable [29]. Integrating synthesis-condition prediction, literature mining and autonomous experimentation with computed stability is a priority [39, 67]. The reproducibility and structural validation of autonomously generated claims also require careful independent confirmation.

Interfaces, defects and operando conditions. Device performance is often controlled by interfaces, grain boundaries, defects and the electrochemical environment rather than by bulk properties. These systems are large, disordered and out of equilibrium, so that they are poorly represented in current training sets. Extending datasets and MLIPs to surfaces, solid–solid interfaces, solvent and applied potential is therefore a key frontier, with the Open Catalyst effort as an early step [37, 38].

Interpretability and physical insight. A purely black-box ranking of candidates provides limited scientific understanding. Combining interpretable models, symbolic regression and physically motivated descriptors with high-capacity networks can lead to design rules that generalize beyond the training data [44, 69]. Hybrid approaches that embed physical constraints, such as symmetry, conservation laws and known scaling relations, into the model architecture are also promising [55].

Data standards and openness. Realizing the full potential of AI requires FAIR (findable, accessible, interoperable, reusable) data, standardized metadata describing calculation settings, and the reporting of negative results from both computation and experiment. Open benchmarks and community challenges accelerate progress by allowing fair comparison of methods [37, 49].

Looking ahead, we expect three developments to be particularly influential. The first is foundation models for atomistic simulation that are pretrained on very large, diverse datasets and fine-tuned for specific energy applications. The second is the tight coupling of generative design with multi-objective optimization, so that voltage, stability, cost and sustainability are considered together. The third is the integration of these computational tools with automated synthesis and characterization to form self-driving laboratories [19, 68].

7. Conclusions

Density functional theory and artificial intelligence are complementary rather than competing approaches. DFT supplies physically grounded, transferable predictions of the stability, electronic structure, transport and reactivity that determine the performance of energy materials, whereas AI supplies the speed, scalability and generative capability that DFT

lacks. Their combination, implemented as a closed-loop workflow of high-throughput calculation, surrogate modeling, inverse design, high-fidelity validation and experimental feedback, has already contributed to the identification of cathode materials, solid electrolytes, perovskite absorbers and electrocatalysts, and has expanded the known landscape of stable inorganic crystals by orders of magnitude.

The remaining obstacles are substantial but well defined: improving the accuracy and consistency of training data, achieving reliable extrapolation with calibrated uncertainty, predicting synthesizability and kinetic behavior, modeling realistic interfaces and operating conditions, and validating predictions experimentally. Progress on these fronts, supported by open data and community benchmarks, will determine how quickly AI–DFT approaches can move from accelerating discovery on the computer to delivering the materials that a sustainable energy system requires.

References

M. Armand and J.-M. Tarascon, Building better batteries, Nature **451**, 652 (2008).

B. Dunn, H. Kamath and J.-M. Tarascon, Electrical energy storage for the grid: A battery of choices, Science **334**, 928 (2011).

J. B. Goodenough and K.-S. Park, The Li-ion rechargeable battery: A perspective, J. Am. Chem. Soc. **135**, 1167 (2013).

P. Simon and Y. Gogotsi, Materials for electrochemical capacitors, Nat. Mater. **7**, 845 (2008).

P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. **136**, B864 (1964).

W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

Aqeel, A., Jabeen, S., Kainat, T., Ahmed, M., Azhar, I. A., Arain, H., & Khanam, Y. (2025). Role of microbial contaminants in the development of malnutrition in pediatric patients with gastrointestinal diseases. Frontier in Medical and Health Research, 3(8), 1215–1225. https://fmhr.net/index.php/fmhr/article/view/1558

P. Giannozzi et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter **21**, 395502 (2009).

Aqeel, A., Mirani, Z. A., Bawa, M. Y., Asif, S., Noor, R., Khan, S., & Abbas, T. (2021). Perceptive on bacteriological quality in foods of animal origin sold in the local market: Potential threats for the perishable food supply chain. International Journal of Endorsing Health Science Research (IJEHSR), Vol. 9 No. 4.

M. K. Aydinol, A. F. Kohan, G. Ceder, K. Cho and J. Joannopoulos, Ab initio study of lithium intercalation in metal oxides and metal dichalcogenides, Phys. Rev. B **56**, 1354 (1997).

Jahid, A., Islam, M. S., & Obaid, M. O. (2024). AI-driven precision oncology prediction: Deep learning–driven early detection of carcinogenesis using multi-omics data integration. Journal of Carcinogenesis, 23(1), 908–915. https://doi.org/10.64149/J.Carcinog.23.1.908-915

A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K. A. Persson, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

S. Curtarolo et al., AFLOW: An automatic framework for high-throughput materials discovery, Comput. Mater. Sci. **58**, 218 (2012).

Saeed., M , Zeb.,I , Saleem., M,I & Usman., S. (2026.)Antimicrobial Resistance Patterns in Clinical Pathogens and Implications for Public Health. Physical Education, Health and Social Sciences.Vol. 4 No. 3. 661-675 https://doi.org/10.63163/jpehss.v4i1.1198

Ishaq, Zeb., I, Saeed., M, Gurganari., L, Parveen., S.(2026). Evaluation of Hematological Changes in COVID-19 Patients and Their Prognostic Value in Clinical Management. Pakistan Journal of Medical & Cardiological Review.556-566

Javed, M. A., Saeed, M., & Zeb, I. (2026). Automated vs Manual Hematology Techniques: Diagnostic Accuracy, Efficiency, and Clinical Utility in Tertiary Care Hospitals: https://doi.org/10.5281/zenodo. 19659917. Pakistan Journal of Medical & Cardiological Review, 5(1), 2771-277

Panhwar, K., Soomro, B. N., Bhatti, S., & Jaskani, F. H. (2025). Novel Encoder–Decoder Architecture with Attention Mechanisms for Satellite-Based Environmental Forecasting in Smart City Applications. Future Internet, 17(9), 380. https://doi.org/10.3390/fi17090380.

O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko and K.-R. Müller, Machine learning force fields, Chem. Rev. **121**, 10142 (2021).

C. Zeni et al., A generative model for inorganic materials design, Nature **639**, 624 (2025).

J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).

J. Sun, A. Ruzsinszky and J. P. Perdew, Strongly constrained and appropriately normed semilocal density functional, Phys. Rev. Lett. **115**, 036402 (2015).

Nazim Nazir. (2023). Artificial intelligence-driven predictive safety instrumented systems for critical energy [23] infrastructure: A framework for next-generation functional safety. International Journal of Communication Networks and Information Security, Volume 15 ,(Issue-12), [pages 18]. DOI-https://ijcnis.org/index.php/ijcnis/article/view/9129

Nazim Nazir. (2022). Intelligent fault diagnosis and resilient control of safety-critical industrial systems using explainable AI. Journal of Information Systems Engineering and Management, Volume 7 ([Issue-4]), [pages-12]. DOI-15_Intelligent_Fault_Diagnosis.pdf

Ashraf, A., Sultani, M. A., Atique, M. S., Qureshi, M., & Shayan, T. J. (2026). Exploring Graphene-Enhanced Quantum Materials and Superconducting Heterostructures for Advanced Batteries and Sustainable Fusion Energy Technologies. Sch Acad J Biosci, 14(8), 555-581. doi.org/10.36347/sajb.2026.v14i08.004

Alam, T., Asif, S., Ayaz, S., Rehman, H., Sana, A., Alam, S., Qamar, F., Naveed, S., & Ghi, A. O. (2019). Total phenolic content and total flavonoid content of Embelia ribes by radical scavenging activity. Latin American Journal of Pharmacy, 38(7), 1467–1471.

A. Jain, G. Hautier, S. P. Ong, C. J. Moore, C. C. Fischer, K. A. Persson and G. Ceder, Formation enthalpies by mixing GGA and GGA+U calculations, Phys. Rev. B **84**, 045115 (2011).

S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. **132**, 154104 (2010).

J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard and H. Jónsson, Origin of the overpotential for oxygen reduction at a fuel-cell cathode, J. Phys. Chem. B **108**, 17886 (2004).

Asif, S., Alam, T., Nawab, A., & Naveed, S. (2016). Simple UV spectrophotometric method development for meropenem in bulk form. RADS Journal of Pharmacy and Pharmaceutical Sciences, 4(1), 45–48.

G. Henkelman, B. P. Uberuaga and H. Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, J. Chem. Phys. **113**, 9901 (2000).

S. P. Ong, V. L. Chevrier, G. Hautier, A. Jain, C. Moore, S. Kim, X. Ma and G. Ceder, Voltage, stability and diffusion barrier differences between sodium-ion and lithium-ion intercalation materials, Energy Environ. Sci. **4**, 3680 (2011).

Y. Mo, S. P. Ong and G. Ceder, First principles study of the $Li_{10}GeP_2S_{12}$ lithium super ionic conductor material, Chem. Mater. **24**, 15 (2012).

Y. Wang, W. D. Richards, S. P. Ong, L. J. Miara, J. C. Kim, Y. Mo and G. Ceder, Design principles for solid-state lithium superionic conductors, Nat. Mater. **14**, 1026 (2015).

W.-J. Yin, T. Shi and Y. Yan, Unusual defect physics in $CH_3NH_3PbI_3$ perovskite solar cell absorber, Appl. Phys. Lett. **104**, 063903 (2014).

A. Fujishima and K. Honda, Electrochemical photolysis of water at a semiconductor electrode, Nature **238**, 37 (1972).

S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, Python Materials Genomics (pymatgen): A robust, open-source Python library for materials analysis, Comput. Mater. Sci. **68**, 314 (2013).

L. Chanussot et al., Open Catalyst 2020 (OC20) dataset and community challenges, ACS Catal. **11**, 6059 (2021).

R. Tran et al., The Open Catalyst 2022 (OC22) dataset and challenges for oxide electrocatalysts, ACS Catal. **13**, 3066 (2023).

V. Tshitoyan, J. Dagdelen, L. Weston, A. Dunn, Z. Rong, O. Kononova, K. A. Persson, G. Ceder and A. Jain, Unsupervised word embeddings capture latent knowledge from materials science literature, Nature **571**, 95 (2019).

G. Pilania, C. Wang, X. Jiang, S. Rajasekaran and R. Ramprasad, Accelerating materials property predictions using machine learning, Sci. Rep. **3**, 2810 (2013).

L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, A general-purpose machine learning framework for predicting properties of inorganic materials, npj Comput. Mater. **2**, 16028 (2016).

R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi and C. Kim, Machine learning in materials informatics: recent applications and prospects, npj Comput. Mater. **3**, 54 (2017).

L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl and M. Scheffler, Big data of materials science: Critical role of the descriptor, Phys. Rev. Lett. **114**, 105503 (2015).

R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler and L. M. Ghiringhelli, SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates, Phys. Rev. Mater. **2**, 083802 (2018).

T. Xie and J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. **120**, 145301 (2018).

C. Chen, W. Ye, Y. Zuo, C. Zheng and S. P. Ong, Graph networks as a universal machine learning framework for molecules and crystals, Chem. Mater. **31**, 3564 (2019).

K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko and K.-R. Müller, SchNet – A deep learning architecture for molecules and materials, J. Chem. Phys. **148**, 241722 (2018).

K. Choudhary and B. DeCost, Atomistic line graph neural network for improved materials property predictions, npj Comput. Mater. **7**, 185 (2021).

V. Fung, J. Zhang, E. Juarez and B. G. Sumpter, Benchmarking graph neural networks for materials chemistry, npj Comput. Mater. **7**, 84 (2021).

C. J. Bartel, A. Trewartha, Q. Wang, A. Dunn, A. Jain and G. Ceder, A critical examination of compound stability predictions from machine-learned formation energies, npj Comput. Mater. **6**, 97 (2020).

W. Ye, C. Chen, Z. Wang, I.-H. Chu and S. P. Ong, Deep neural networks for accurate predictions of crystal stability, Nat. Commun. **9**, 3800 (2018).

J. Behler and M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, Phys. Rev. Lett. **98**, 146401 (2007).

A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, Phys. Rev. Lett. **104**, 136403 (2010).

V. L. Deringer, M. A. Caro and G. Csányi, Machine learning interatomic potentials as emerging tools for materials science, Adv. Mater. **31**, 1902765 (2019).

S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt and B. Kozinsky, E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, Nat. Commun. **13**, 2453 (2022).

I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner and G. Csányi, MACE: Higher order equivariant message passing neural networks for fast and accurate force fields, Adv. Neural Inf. Process. Syst. **35** (2022).

C. Chen and S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, Nat. Comput. Sci. **2**, 718 (2022).

B. Deng, P. Zhong, K. Jun, J. Riebesell, K. Han, C. J. Bartel and G. Ceder, CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling, Nat. Mach. Intell. **5**, 1031 (2023).

A. Zunger, Inverse design in search of materials with target functionalities, Nat. Rev. Chem. **2**, 0121 (2018).

B. Sanchez-Lengeling and A. Aspuru-Guzik, Inverse molecular design using machine learning: Generative models for matter engineering, Science **361**, 360 (2018).

T. Xie, X. Fu, O.-E. Ganea, R. Barzilay and T. Jaakkola, Crystal diffusion variational autoencoder for periodic material generation, Int. Conf. Learn. Represent. (ICLR) (2022).

A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon and E. D. Cubuk, Scaling deep learning for materials discovery, Nature **624**, 80 (2023).

C. J. Pickard and R. J. Needs, Ab initio random structure searching, J. Phys.: Condens. Matter **23**, 053201 (2011).

A. R. Oganov and C. W. Glass, Crystal structure prediction using ab initio evolutionary techniques: Principles and applications, J. Chem. Phys. **124**, 244704 (2006).

B. Shahriari, K. Swersky, Z. Wang, R. P. Adams and N. de Freitas, Taking the human out of the loop: A review of Bayesian optimization, Proc. IEEE **104**, 148 (2016).

T. Lookman, P. V. Balachandran, D. Xue and R. Yuan, Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design, npj Comput. Mater. **5**, 21 (2019).

N. J. Szymanski et al., An autonomous laboratory for the accelerated synthesis of novel inorganic materials, Nature **624**, 86 (2023).

D. P. Tabor et al., Accelerating the discovery of materials for clean energy in the era of smart automation, Nat. Rev. Mater. **3**, 5 (2018).

S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. **30** (2017).

G. Ceder, Opportunities and challenges for first-principles materials design and applications to Li battery materials, MRS Bull. **35**, 693 (2010).

G. Hautier, A. Jain, H. Chen, C. Moore, S. P. Ong and G. Ceder, Phosphates as lithium-ion battery cathodes: An evaluation based on high-throughput ab initio calculations, Chem. Mater. **23**, 3495 (2011).

A. Urban, D.-H. Seo and G. Ceder, Computational understanding of Li-ion batteries, npj Comput. Mater. **2**, 16002 (2016).

C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng and S. P. Ong, A critical review of machine learning of energy materials, Adv. Energy Mater. **10**, 1903242 (2020).

N. Kamaya et al., A lithium superionic conductor, Nat. Mater. **10**, 682 (2011).

A. D. Sendek, Q. Yang, E. D. Cubuk, K.-A. N. Duerloo, Y. Cui and E. J. Reed, Holistic computational structure screening of more than 12000 candidates for solid lithium-ion conductor materials, Energy Environ. Sci. **10**, 306 (2017).

W. Shockley and H. J. Queisser, Detailed balance limit of efficiency of p-n junction solar cells, J. Appl. Phys. **32**, 510 (1961).

A. Kojima, K. Teshima, Y. Shirai and T. Miyasaka, Organometal halide perovskites as visible-light sensitizers for photovoltaic cells, J. Am. Chem. Soc. **131**, 6050 (2009).

A. Walsh and A. Zunger, Instilling defect tolerance in new compounds, Nat. Mater. **16**, 964 (2017).

S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li and J. Wang, Accelerated discovery of stable lead-free hybrid organic-inorganic perovskites via machine learning, Nat. Commun. **9**, 3405 (2018).

B. Hammer and J. K. Nørskov, Why gold is the noblest of all the metals, Nature **376**, 238 (1995).

J. Greeley, T. F. Jaramillo, J. Bonde, I. B. Chorkendorff and J. K. Nørskov, Computational high-throughput screening of electrocatalytic materials for hydrogen evolution, Nat. Mater. **5**, 909 (2006).

J. K. Nørskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, Towards the computational design of solid catalysts, Nat. Chem. **1**, 37 (2009).

I. C. Man et al., Universality in oxygen evolution electrocatalysis on oxide surfaces, ChemCatChem **3**, 1159 (2011).

J. Rossmeisl, Z.-W. Qu, H. Zhu, G.-J. Kroes and J. K. Nørskov, Electrolysis of water on oxide surfaces, J. Electroanal. Chem. **607**, 83 (2007).

A. A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl and J. K. Nørskov, How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels, Energy Environ. Sci. **3**, 1311 (2010).

M. Zhong et al., Accelerated discovery of CO_2 electrocatalysts using active machine learning, Nature **581**, 178 (2020).

Z. W. Seh, J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov and T. F. Jaramillo, Combining theory and experiment in electrocatalysis: Insights into materials design, Science **355**, eaad4998 (2017).