

Review

Open Access



AI in single-atom catalysts: a review of design and applications

Qiumei Yu¹, Ninggui Ma^{1,2,*}, Chihon Leung², Han Liu^{2,3}, Yang Ren^{2,3} , Zhanhua Wei^{1,*}

¹Xiamen Key Laboratory of Optoelectronic Materials and Advanced Manufacturing, Institute of Luminescent Materials and Information Displays, College of Materials Science and Engineering, Huaqiao University, Xiamen 361000, Fujian, China.

²Department of Physics, City University of Hong Kong, Hong Kong 999077, China.

³Shenzhen Research Institute, City University of Hong Kong, Shenzhen 518000, Guangdong, China.

*Correspondence to: Prof. Ninggui Ma, Prof. Zhanhua Wei, Xiamen Key Laboratory of Optoelectronic Materials and Advanced Manufacturing, Institute of Luminescent Materials and Information Displays, College of Materials Science and Engineering, Huaqiao University, 668 Jimei Avenue, Jimei District, Xiamen 361000, Fujian, China. E-mail: maninggui@163.com; weizhanhua@hqu.edu.cn

How to cite this article: Yu, Q.; Ma, N.; Leung, C.; Liu, H.; Ren, Y.; Wei, Z. AI in single-atom catalysts: a review of design and applications. *J. Mater. Inf.* 2025, 5, 9. <https://dx.doi.org/10.20517/jmi.2024.78>

Received: 27 Nov 2024 **First Decision:** 19 Dec 2024 **Revised:** 15 Jan 2025 **Accepted:** 22 Jan 2025 **Published:** 12 Feb 2025

Academic Editors: Hao Li, Yaqiong Su **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

Single-atom catalysts (SACs) have emerged as a research frontier in catalytic materials, distinguished by their unique atom-level dispersion, which significantly enhances catalytic activity, selectivity, and stability. SACs demonstrate substantial promise in electrocatalysis applications, such as fuel cells, CO₂ reduction, and hydrogen production, due to their ability to maximize utilization of active sites. However, the development of efficient and stable SACs involves intricate design and screening processes. In this work, artificial intelligence (AI), particularly machine learning (ML) and neural networks (NNs), offers powerful tools for accelerating the discovery and optimization of SACs. This review systematically discusses the application of AI technologies in SACs development through four key stages: (1) Density functional theory (DFT) and *ab initio* molecular dynamics (AIMD) simulations: DFT and AIMD are used to investigate catalytic mechanisms, with high-throughput applications significantly expanding accessible datasets; (2) Regression models: ML regression models identify key features that influence catalytic performance, streamlining the selection of promising materials; (3) NNs: NNs expedite the screening of known structural models, facilitating rapid assessment of catalytic potential; (4) Generative adversarial networks (GANs): GANs enable the prediction and design of novel high-performance catalysts tailored to specific requirements. This work provides a comprehensive overview of the current status of AI applications in SACs and offers insights and recommendations for future advancements in the field.

Keywords: Single-atom catalysts, AI, machine learning



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



INTRODUCTION

With the rapid advancement of technology, energy crisis and environmental pollution have emerged as two core challenges faced by societies worldwide. Rising energy demand and depleting fossil fuels intensify energy crisis, while industrialization and human activities aggravate pollution. Effectively addressing these two issues has become a central focus for the global scientific and industrial communities. Against this backdrop, catalysts have gained prominence as a key technology in tackling energy and environmental challenges, due to their unique advantages in accelerating chemical reactions, improving energy conversion efficiency, and reducing pollutant emissions. For instance, by enhancing energy conversion efficiency, catalysts can reduce energy consumption. Platinum-based catalysts demonstrate excellent catalytic activity for oxygen reduction reaction (ORR) in fuel cells^[1-6]. Iron-based single-atom catalysts (SACs) have been found to significantly reduce energy requirements for water-splitting reactions, greatly improving efficiency and cost-effectiveness of hydrogen production^[7-9]. Hydrogen production technology is a crucial avenue for future clean energy development, and efficient catalysts play an indispensable role in this process^[10]. Moreover, catalysts are widely applied in the development and utilization of renewable energy sources. For example, titanium-based catalysts can efficiently catalyze the photocatalytic water-splitting reaction, converting solar energy into hydrogen and providing a sustainable energy solution^[11,12]. Furthermore, catalysts significantly enhance the selectivity and yield of reactions in biomass conversion processes, facilitating the efficient utilization of biomass resources and providing essential technological support for the development of biomass energy^[13].

In environmental pollution control, catalysts also play a crucial role. For instance, in automotive exhaust treatment, they convert harmful gases such as carbon monoxide and nitrogen oxides into harmless substances, thereby reducing emissions of atmospheric pollutants^[14,15]. Additionally, catalysts have widespread applications in wastewater treatment^[16,17], pollutant degradation, and carbon dioxide reduction^[18,19]. Thus, catalysts not only enhance energy conversion efficiency but also play a critical role in reducing pollutant emissions and improving environmental quality.

However, the current research and development of catalysts still confront several challenges. Firstly, many efficient catalysts rely on precious metal materials, such as platinum and palladium. The high cost and limited natural availability of these metals restrict widespread industrial application^[14,20,21]. Secondly, some catalysts exhibit relatively low catalytic efficiency, leading to challenges in effectively treating harmful pollutants. These issues have prompted researchers to seek alternative solutions that are both efficient and cost-effective^[22,23].

Against this backdrop, SACs have garnered significant attention as an emerging technology^[24]. They are specialized supported metal catalysts with active metal components dispersed as single atoms. By precisely designing the chemical coordination environment around single atoms, researchers can customize active sites^[25,26] to exhibit unique catalytic properties. This design strategy not only provides a novel research platform for catalytic reactions but also offers significant opportunities for exploring and understanding catalytic mechanisms. Despite the immense potential of SACs, their development still faces challenges. The preparation of SACs involves multiple complex steps, including material design, precise control of active sites, and optimization of catalytic performance^[27-29]. Traditional experimental methods often rely on experience and trial-and-error approaches, which can be inefficient and may not achieve optimal performance. In recent years, advances in density functional theory (DFT) calculations have significantly accelerated the screening of efficient catalysts^[30-35]. However, these calculations require substantial computational resources, resulting in significant consumption of CPU and GPU power, along with high

time costs^[36]. Based on this background, artificial intelligence (AI), particularly machine learning (ML) and neural networks (NNs), has revolutionized materials science by providing novel methodologies for the discovery and design of SACs^[37-40].

By analyzing large datasets from experimental results and theoretical calculations, AI can extract potential key parameters and establish multi-layered predictive models to forecast catalyst performance under various reaction conditions^[41-44]. This data-driven approach can significantly accelerate the screening and optimization process for catalysts. For example, ML models can identify the optimal catalyst combinations within hours, greatly reducing time and resource consumption. Additionally, AI can uncover the intrinsic relationships between catalyst structure and performance, providing theoretical support for further catalyst optimization^[45]. It also can predict catalyst performance under novel reaction conditions, enabling preliminary assessments of catalyst effectiveness and adjustment of experimental designs in the early stages. AI-driven automated data analysis and model optimization techniques facilitate rapid iterations of experimental designs by simulating catalyst performance under various reaction conditions, ultimately identifying catalyst solutions with high activity and stability^[46-50].

All in all, this article aims to comprehensively review the latest advancements of AI technologies in the research of SACs and explore their potential applications in future materials science. By leveraging the strengths of AI technologies, significant breakthroughs are anticipated in the development of efficient and stable catalysts, providing crucial support in addressing global energy and environmental challenges. The article analyzes four key stages of AI in the field of electrochemical catalysts [Figure 1]: firstly, generating data through DFT and high-throughput screening (HTS) to create databases suitable for AI processing; secondly, using ML regression models to analyze data and conduct feature importance analysis to identify key characteristics affecting catalyst; thirdly, applying NNs to rapidly screen candidates with potential high catalytic activity; and finally, employing generative adversarial networks (GANs) to design efficient catalysts that meet specific requirements.

GENERATE DATABASE UTILIZING HTS IN COMBINATION WITH DFT AND *AB INITIO* MOLECULAR DYNAMICS

The initial stage: the DFT and *ab initio* molecular dynamics are calculated to screen the performance of the catalyst one by one

In the early stages of developing SACs, researchers primarily relied on DFT and *ab initio* molecular dynamics (AIMD). DFT is widely used to predict the stability, electronic structure, reaction pathways, and energy barriers of SACs^[51,52]. The catalytic activity of SACs is closely linked to the electronic structure of individual active atoms and their interactions with the support, making DFT calculations essential for simulating binding energies, adsorption energies, and reaction transition states. For example, the adsorption behavior of metal atoms on different supports, such as graphene, nitrogen-doped carbon, and titanium dioxide, can be systematically studied using DFT^[53,54]. Complementarily, AIMD integrates with DFT, simulating the spatiotemporal evolution of materials at the atomic scale to investigate the thermodynamic behavior of catalysts under realistic reaction conditions. AIMD not only validates the accuracy of DFT-optimized models but also simulates the kinetic stability of catalysts under varying temperatures and pressures, providing dynamic information crucial for predicting catalyst performance in complex reaction environments.

These computational methods can generate a wealth of structural and performance data. On the one hand, they serve as valuable empirical information; on the other hand, they provide a crucial foundation for new research. Based on existing physicochemical information, researchers can make informed predictions about

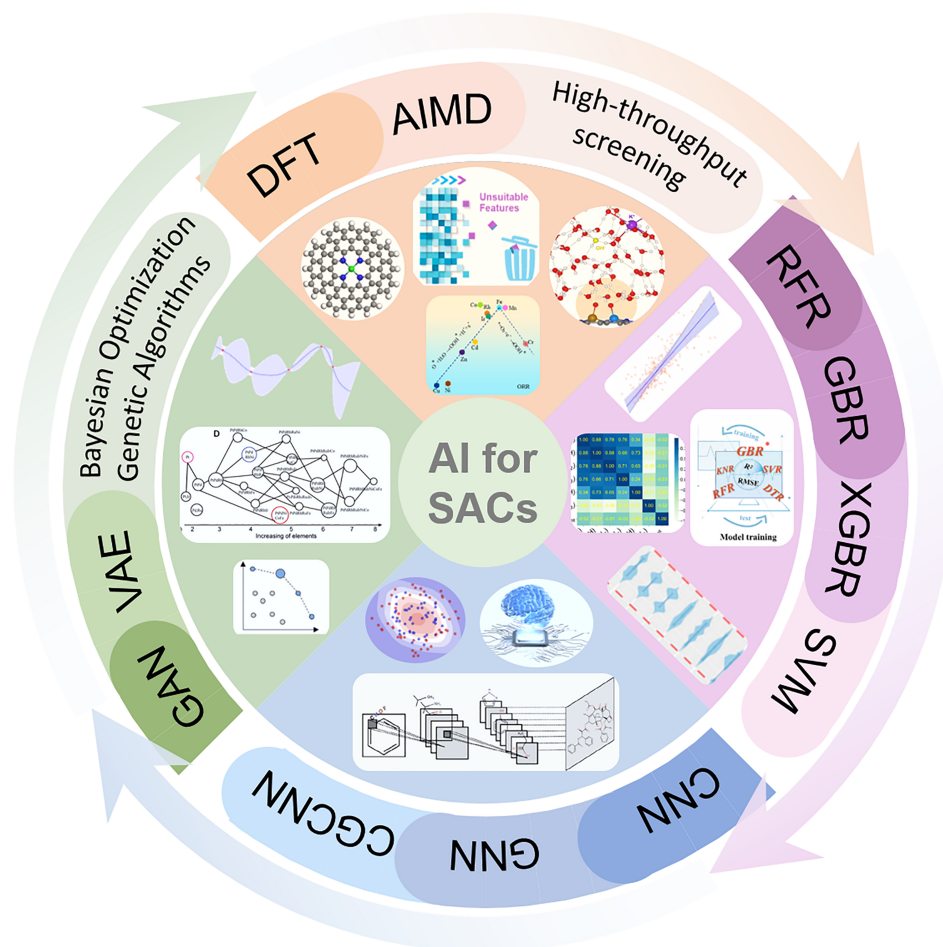


Figure 1. Advancements and implementations of AI in SACs: facilitating HTS through DFT database construction, accelerating feature importance analysis via ML for enhanced screening, crystal structure analysis utilizing NN, and generation of potential atomic structure models through GANs. AI: Artificial intelligence; SACs: single-atom catalysts; HTS: high-throughput screening; DFT: density functional theory; ML: machine learning; NN: neural network.

the performance of similar materials and subsequently validate these predictions through experiments. This integration of theory and experimentation has driven deeper research and technological advancements in the field of SACs.

The necessity of DFT and AIMD in SACs research

SACs demonstrate tremendous application potential in the energy and environmental sectors due to their unique catalytic properties. However, traditional *d*-band theory struggles to explain the catalytic mechanisms of SACs, presenting ongoing challenges in theoretical characterization^[55]. To address this issue, researchers have made significant progress through the synergistic use of DFT and AIMD.

DFT can be employed to investigate the fundamental mechanisms of molecular adsorption in SACs, thus unveiling the reaction pathways. For example, He *et al.* investigated the catalytic oxidation of CO on SACs using DFT calculations, proposing a “selective orbital coupling” mechanism that reveals the selective coupling between the localized *d* orbitals of metal single atoms and the π^* orbitals of O₂ molecules^[56]. This coupling determines the strength of the M–O bond and explains the variations in energy barriers along the

reaction pathway. By calculating the energies of different orbitals, they quantitatively predicted the adsorption strength and correlated it with the reaction barriers, enhancing the understanding of catalytic behavior of SACs. Using DFT calculations, researchers analyzed the highly localized d orbital characteristics of metal doping. The Wannier functions for the coupling between the d and π^* orbitals of O_2 molecules in various adsorption configurations [Figure 2A] illustrate that the most stable structure of O_2 molecules arises from the selective coupling of the π^* orbitals with specific d orbitals ($d_{x^2-y^2}$) of metal dopants, forming the M–O bond. This coupling mechanism determines the strength of the M–O bond and explains the trends in reaction energy barriers during the CO oxidation reaction. Figure 2B quantitatively relates adsorption strength to reaction barriers by utilizing the energy arrangement of the selected orbitals, helping to understand how the chosen d states influence reaction behavior. DFT calculations revealed the electronic structure characteristics of SACs, providing a basis for understanding catalytic mechanisms. Our previous research also demonstrated that the sub-orbitals of d electrons play different roles in the adsorption of catalytic intermediates. For example, when surface transition metals (TM) bind with adsorbed molecules, the coupling of the $3\sigma \rightarrow d_{z^2}$ orbital primarily exhibits an attractive effect, while the $d_{xz}/d_{yz} \rightarrow 2\pi^*$ coupling mainly displays a repulsive effect [Figure 2C and D]. Consequently, Ti_2CO_2 shows anomalous d -band adsorption characteristics for *N , *NH , *NH_2 , and *NH_3 under strain^[57]. A similar phenomenon was also observed during stretching or compressing $Pt-V_F-Ti_2CF_2$, resulting in a paradoxical situation where the longer Pt–O bond length in *OH exhibited a weaker adsorption energy^[58].

Due to incomplete mechanistic understanding of the ORR on Fe–N–C material systems, Hutchison *et al.* have reported a fifth-coordinate structure for H_2O formation during ORR on $Fe-N_4-C$, based on a combined study employing DFT and AIMD^[59]. Under potentials relevant to the ORR, OH is converted to H_2O . The results indicated that the Fe(III/II) oxidation-reduction potentials and the ORR onset potentials closely resemble experimental findings. Reliable predictions of ORR onset potential and Fe(III/II) oxidation-reduction potentials are achieved when $Fe^{III}-OH$ converts to Fe^{II} , and the desorption of H_2O necessitates axial co-adsorption of H_2O onto the iron center. Considering that the five-coordinate model spontaneously forms and exhibits ORR chemistry consistent with experimental measurements in AIMD simulations, it serves as the fundamental structure for simulating the chemistry of Fe–N–C active sites in experimentally relevant systems [Figure 2E].

Additionally, Xiao *et al.* employed a combined approach of DFT and AIMD to systematically investigate the catalytic performance of nitrogen-coordinated TM carbon materials (TM– N_x –C) in the ORR and the oxygen evolution reaction (OER)^[60]. They optimized the geometric models of TM– N_3 –C and TM– N_4 –C using DFT and calculated the formation and binding energies for multiple TM atoms. Ultimately, they selected seven chemically stable metals (Mn, Fe, Co, Ni, Cu, Pd, Pt) for detailed analysis [Figure 2F]. The results from DFT calculations indicated that Ni– N_3 –C exhibited the best catalytic performance in the ORR/OER, featuring optimal adsorption energy and the lowest overpotential [Figure 2G–I].

To further validate the reliability of these models, AIMD was used to assess the thermodynamic stability of Ni– N_3 –C and Pd– N_3 –C at different temperatures. The results shown in Figure 3A and B indicated that Ni– N_3 –C remains stable at high temperatures, while Pd– N_3 –C is stable only below 400 K, further confirming the potential of Ni– N_3 –C as an efficient bifunctional oxygen catalyst.

The aforementioned studies provide strong evidence that DFT and AIMD play crucial roles in the field of catalysis. They serve an irreplaceable function in advancing catalytic theory.

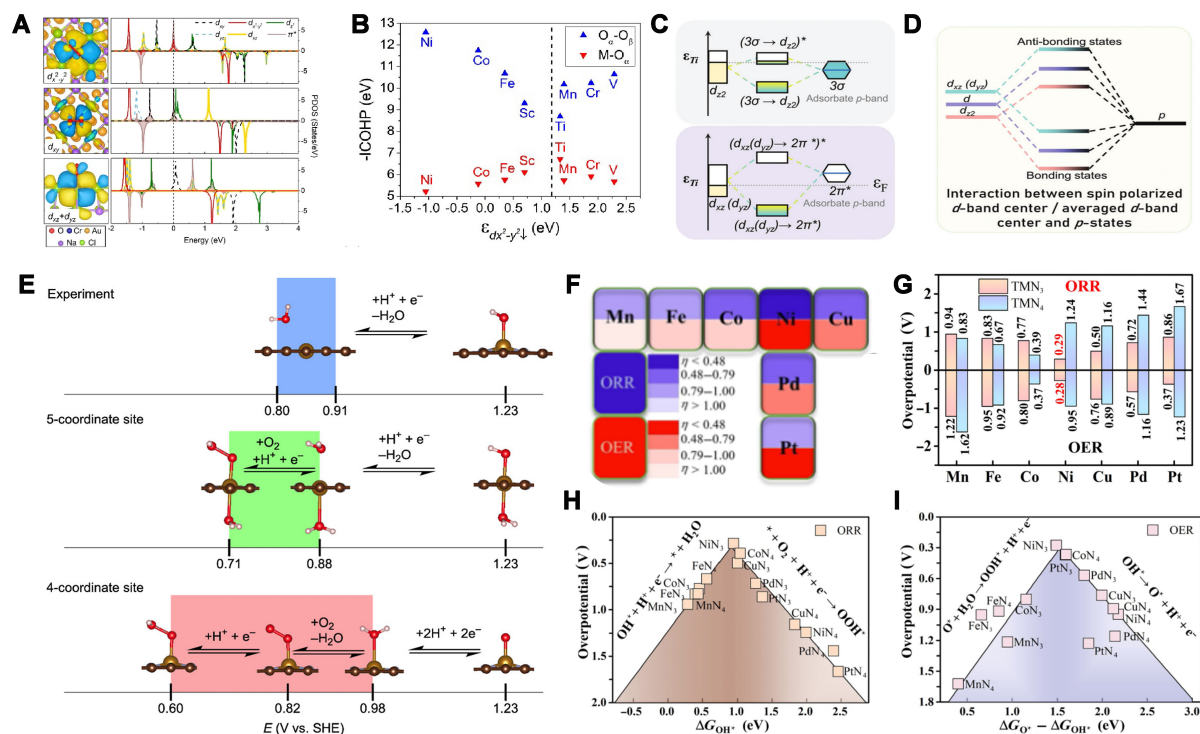


Figure 2. (A) Wannier functions of d orbitals and π orbital coupling in O_2 molecules with various adsorption configurations; (B) Correlation between adsorption strength and reaction barrier, with variation of $-ICOHP$ for $M-O_\alpha$ and $O_\alpha-O_\beta$ bonds as a function of $d_{x^2-y^2}$. Copyright 2024, American Chemical Society, Reproduced with permission^[56]; (C) Mechanism of “donation and backdonation” in N_2 adsorption on titanium: interaction between d_{z^2} orbital of Ti and 3σ orbital of N_2 , and between d_{xz}/d_{yz} orbitals of Ti and $2\pi^*$ orbitals of N_2 ; (D) Coupling interactions between adsorbate states and both the average d -band center and spin-polarized d -band center: the complexity arising from d -orbital splitting and its influence on various adsorption behaviors. Copyright 2023, John Wiley and Sons, Reproduced with permission^[57]; (E) Illustration of ORR cycle: a comparison between experimental ORR mechanism and explored five-coordinate model, with the five-coordinate model predicting a closer starting potential to experimental results. Copyright 2022, American Chemical Society, Reproduced with permission^[59]; (F) Illustrative diagram of TM considered in this study, with distinct colors representing corresponding ranges of ORR/OER overpotentials (η); (g) Potential overvoltages for ORR and OER across all TM- N_x -C catalysts; Volcano plots for ORR (H) and OER (I) on the TM- N_x -C system. Copyright 2021, Springer Nature, Reproduced with permission^[60]. ORR: Oxygen reduction reaction; TM: transition metals; OER: oxygen evolution reaction.

The widespread use of DFT and AIMD in SACs

Graphdiyne, as an emerging two-dimensional carbon material, possesses unique electronic structures and exhibits highly efficient catalytic reduction properties^[61]. Based on DFT, Ullah *et al.* designed and studied the application of $3d$ TM SACs for the hydrogen evolution reaction (HER) on graphyne (GY) surfaces^[62]. By optimizing the geometric structure of the TM@GY composite and conducting frequency analysis, they confirmed the minimal energy state of the structure. Among all the systems considered, the nickel SAC anchored to the GY support demonstrated the highest thermodynamic stability. The adsorption energies and hydrogen adsorption free energy (ΔG_H) were calculated to evaluate the hydrogen adsorption performance and HER catalytic activity of the SACs. The results indicated that Ni@GY possesses the lowest hydrogen adsorption energy and surface ΔG_H change, showcasing the best HER performance [Figure 3C]. This catalyst represents an excellent SAC with outstanding HER catalytic activity and stability, while also being abundant and cost-effective. Overall, this research offers a theoretical basis for the design of low-cost, high-performance HER catalysts, particularly in the context of integrating TM on two-dimensional carbon materials.

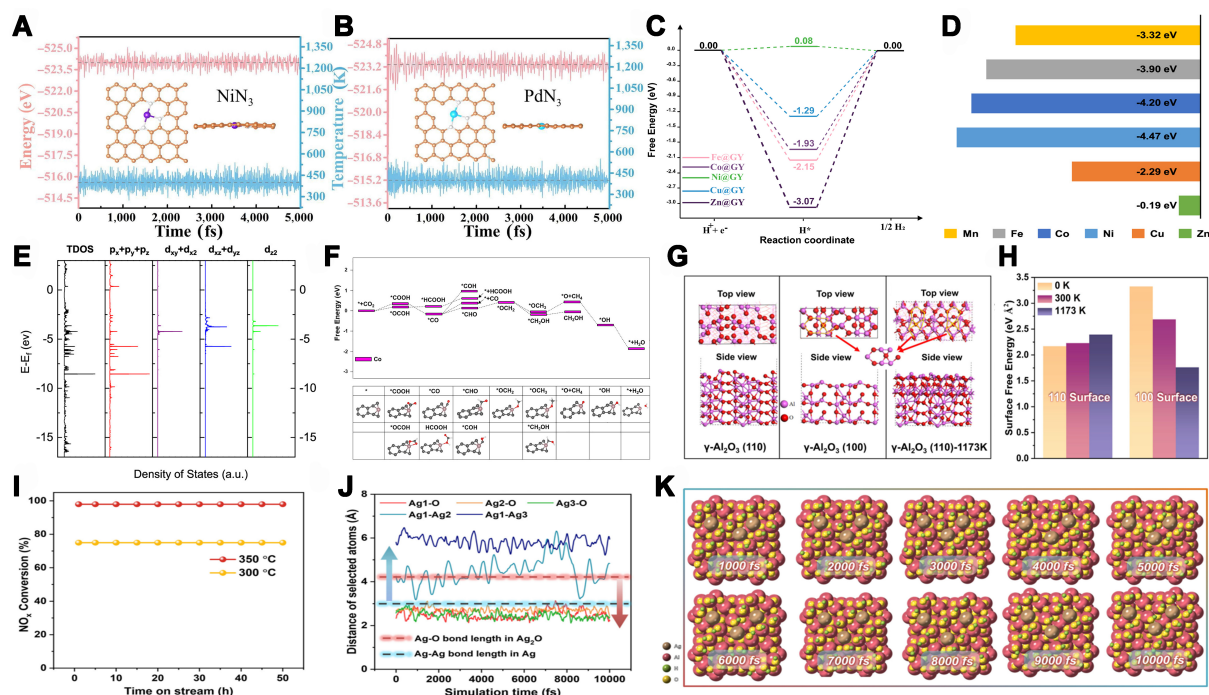


Figure 3. (A and B) Energy and temperature variations of two catalysts over 5 ps at 400 K vs. the AIMD simulation time. Copyright 2021, Springer Nature, Reproduced with permission^[60]; (C) Free energy diagram of HER over TM@GY SACs. Copyright 2021, Elsevier, Reproduced with permission^[62]; (D) Binding energies of various dopant atoms (Zn, Ni, Mn, Fe, Cu, and Co); (E) DOS for Cu-graphdiyne; (F) Gibbs free energy changes of Co-graphdiyne along the CO₂RR pathway. Copyright 2021, American Chemical Society, Reproduced with permission^[63]; (g) γ -Al₂O₃ (110) and γ -Al₂O₃ (100) crystal surface structures; (H) Surface free energy of γ -Al₂O₃ (110) and (100) crystal planes at different temperatures; (I) Stability test of Ag/(Al-900) sample; (J) Time dependence of bond lengths (Å) during the AIMD simulation (10,000 fs); (K) Snapshot of single Ag atom on the γ -Al₂O₃ (100) surface at different simulation times. Copyright 2024, Springer Nature, Reproduced with permission^[64]. AIMD: *Ab initio* molecular dynamics; HER: hydrogen evolution reaction; SACs: single-atom catalysts; DOS: density of states.

Liu *et al.* have also explored the CO₂ reduction reaction (CO₂RR) utilizing 3d TM SACs supported on graphdiyne^[63]. Through DFT calculations, they optimized the geometric structures of single atoms (Mn, Fe, Co, Ni, Cu, and Zn) on the GY support, calculating the binding energies of each doped atom [Figure 3D], the density of states (DOS) with orbital hybridization [Figure 3E], and the Gibbs free energy changes of the intermediates along the CO₂RR pathway [Figure 3F]. This analysis allowed the identification of the rate-limiting steps (RDS) and energy barriers, ultimately reporting several high-activity CO₂RR catalysts. The results indicated that the *sp* hybridization between carbon and metal plays a crucial role in modulating catalytic activity. Additionally, Kan *et al.* systematically investigated the dissociation energy of O₂ when Pt is supported on different MXene surfaces, determining that the ORR on MXene-Pt doping proceeds primarily via a four-electron coupled mechanism rather than a dissociation mechanism^[30]. Further research revealed that high coverage of Pt does not enhance the catalyst’s activity, providing theoretical validation for the rational design and feasibility of SACs^[31].

Using AIMD can reveal the influence of temperature on the phase transition of γ -Al₂O₃ and the anchoring of Ag. In the study of Li *et al.*, understanding the “terminal hydroxyl anchoring mechanism” of Ag on the Al₂O₃ surface is essential for optimizing the state of Ag as an active species and enhancing catalytic performance^[64]. DFT was employed to investigate the surface hydroxyl content of γ -Al₂O₃ (100) and (110) facets. It was found that the (100) facet has a higher density of terminal hydroxyls [Figure 3G], providing more anchoring sites for the dispersion of Ag atoms. By calculating the surface free energies of γ -Al₂O₃ (110)

and (100) facets at different temperatures [Figure 3H], it was observed that the surface free energy of the (110) facet increases at high temperatures, while that of the (100) facet decreases, explaining the phase transition due to high-temperature calcination.

AIMD simulations of the γ -Al₂O₃ (110) facet at high temperatures showed that elevated temperatures lead to the rearrangement of surface atoms, forming a structure similar to the (100) facet and increasing the terminal hydroxyl content. Stability simulations of Ag single atoms on the γ -Al₂O₃ (100) facet [Figure 3I] indicated that it could stably anchor to the (100) surface without aggregation, explaining the dispersion of Ag single atoms in the Ag/(Al-900) sample. Further AIMD simulations of the thermal stability of Ag single atoms on the γ -Al₂O₃ (100) facet [Figure 3J and K] demonstrated that Ag single atoms can remain stable for 10,000 s at 773 K, indicating that the Ag/(Al-900) sample exhibits good thermal stability.

Additionally, Wang *et al.* at Tsinghua University reported the oxidation mechanism of CO at the interface through static DFT and AIMD simulations^[65]. They constructed Au clusters on the surface of CeO₂, where a single Au⁺ atom on the cluster surface acts as an electron acceptor during the reaction, enhancing the adsorption and transport of CO. This significantly lowers the barrier for the reduction of CeO₂ and effectively promotes the CO oxidation reaction. Interestingly, the Au⁺-CO ion appears only in the presence of CO. Once CO is removed, Au⁺ recombines with the Au nanoparticles. This single atom effectively couples the redox process with that of the support, thereby enhancing the overall redox activity, while the Au nanoparticles show little evidence of coupling with the oxidation state of the oxide during the catalytic cycle. This study clarifies an important concept: the actual catalytic active sites may only manifest during the reaction, becoming hidden before and after the process. Thus, the formation of active sites is a dynamic process occurring at the interface between the supported oxide and the metal particles.

Fan *et al.* from Xiamen University believe that sub-nanometer metal clusters in catalysts possess numerous metastable structures, which can interconvert during catalytic reactions, leading to complex catalytic behaviors. Furthermore, comparing the diffusion energy barriers calculated by DFT and AIMD reveals that the static energy barriers from DFT calculations are higher than those from dynamic reactions. Further studies indicate that the formation of Cu₃O increases the melting temperature of the clusters, resulting in a decrease in the entropy of the dissociation products^[66]. This work demonstrates the significant impact of surface adsorption on the dynamic phase transition behavior of clusters and provides a new perspective on dynamic catalysis.

By synergistically employing DFT and AIMD, researchers can deeply assess the catalytic performance and thermodynamic stability of SACs from both static electronic structures and dynamic reaction behaviors. However, DFT and AIMD still possess certain limitations, such as system size constraints and increased computational costs for large systems. The computational cost of AIMD is a barrier for large-scale, long-duration simulations, and it struggles to capture rare microscopic events.

Advanced simulation tools and methods for enhancing DFT calculations

In laboratory research, activity and selectivity are commonly used parameters to evaluate the performance of SACs, while the importance of stability is often overlooked^[67]. SACs exhibit exceptional atomic efficiency and catalytic performance, yet stability remains a significant challenge^[25,68]. The intricate relationship between structure and stability is seldom explored due to degradation complexity and reaction conditions. To achieve more accurate electrochemical simulations, it is essential to consider the environmental impact, including solvents, pH, and electrical potentials.

The Pourbaix diagram directly indicates system stability under specific potential and pH. Di Liberto *et al.* utilized DFT and Pourbaix diagrams to anticipate SAC stability across various pH and voltage ranges^[69]. By integrating experimental data with DFT, the stability of four TM atoms such as Cr, Mn, Fe, and Co dropped on three carbon-based supports was examined. DFT was instrumental in calculating binding energies and Gibbs free energies, and constructing Pourbaix diagrams that visualize SACs stability under varied conditions. It was discovered that under operational conditions, many potential catalysts may dissolve or transform, particularly under oxidizing conditions.

Traditional DFT calculations are conducted under the constant charge model (CCM), whereas practical electrochemical reactions occur under the constant potential model (CPM). Tan *et al.* compared hydrogen adsorption of metal SACs on graphene (M-NC) under both models by employing DFT and Grand Canonical DFT methods, contrasting CCM with CPM^[70]. CCM neglects the influence of the electrode potential, causing deviations in the calculation of $\Delta G(^H)$. In contrast, CPM provides a more accurate depiction of electrocatalytic conditions, and is vital for evaluating HER activity of M-NC SACs.

Cui *et al.* also established a structure-stability relationship for N-doped carbon-supported SACs under CO₂ reduction conditions through advanced CPM and DFT^[71]. Using CP-VASP (a patch to the Vienna Ab-Initio Simulation Package) code, they simulated actual CO₂RR operations, accounting for pH and potential, thereby rendering the computational results closer to real-world scenarios. The study systematically analyzed various factors influencing stability and highlighted metal atom leaching as a critical concern. Strategies for enhancing stability were further experimentally validated.

These findings fill the current stability knowledge gap of SACs under practical operating conditions and are expected to propel their widespread application in sustainable energy systems.

Micro-kinetics is crucial for unraveling catalytic mechanisms and kinetics. Implicit and explicit solvation models can address solvent effects^[72]. Implicit model using a polarizable medium while constructing an electric field to describe the charge distribution of solvent. In contrast, explicit model precisely incorporates solvent molecules, atoms, and cations into the computational system, allowing for direct observation of interactions. Zhang and Li conducted large-scale sampling and investigated the point of zero charge (PZC) and solvation effects of M-N-C catalysts, finding explicit models that can offer more precise predictions^[73]. Incorporating PZC and solvation effects into micro-kinetic models could be considered in future studies to enhance prediction accuracy.

Fe-N-C materials are promising for ORR catalysis, but pH-dependent activity and origins have been a development hurdle. Liu *et al.* unraveled the pH-dependent mechanism in Fe-N-C materials through first principle and micro-kinetic^[74]. By considering the effects of pH, solvation, and electrode potential in micro-kinetic simulations, it was found that the FeN₄ centers are covered by ^{*}OH and ^{*}O intermediates in acidic and alkaline media, respectively. The ^{*}O intermediate optimizes the electronic structure more effectively than the ^{*}OH intermediate, leading to higher ORR activity in alkaline media. Micro-kinetic model was employed to simulate polarization curves and Tafel slopes, with results consistent with laboratory observations. This work provides a quantitative description for understanding ORR performance of Fe-N-C catalysts.

Machine-learned force fields (MLFFs) enhance simulation efficiency and accuracy, overcoming traditional force field limitations. AIMD accelerated by MLFFs can reduce simulation times dramatically, aiding in predicting microscopic dynamics^[75,76]. Zhang *et al.* have integrated MLFFs into molecular dynamics,

overcoming limitations in accuracy and timescale^[77]. Their work has shown that MLFFs can achieve high accuracy while improving computational efficiency, but further enhancements are needed, particularly in high-quality data generation.

In this section, an overview of novel methodologies, including Pourbaix diagrams, CPM, micro-kinetics, and MLFFs, in materials computing highlights potential in advancing understanding and application of SACs in sustainable energy systems.

High-throughput computational acquisition data

Compared to traditional experimental exploration methods, DFT significantly enhances the efficiency of catalyst screening but is limited to single calculations. In the bulk screening of efficient catalysts, substantial human effort is still required to organize and submit computational tasks, often leading to resource idling and waste. Therefore, reducing labor input and simplifying repetitive operations has become a critical issue.

High-throughput computing effectively addresses this challenge. By enabling batch submissions of computational tasks, HTS accelerates the screening process as an efficient method for data collection and processing. When combined with DFT and AIMD results, HTS can systematically gather structural and performance data of various SACs, creating a comprehensive database^[78]. Additionally, HTS allows for the rapid selection of samples with target structures from large databases, systematically evaluating catalyst performance. This method can identify and eliminate inefficient catalysts in the early stages, expediting the design and discovery of efficient SACs. By narrowing the experimental scope, HTS not only saves time and costs but also significantly improves the efficiency of identifying effective catalysts.

Comparative studies on the catalytic performance of carbon-based SACs are relatively scarce. To efficiently identify catalysts for the ORR from a multitude of candidates, HTS is a commonly used and effective approach^[79]. Researchers established a database of 48 candidate SACs composed of six TM elements and eight carbon-based supports. Using DFT, they conducted HTS on 180 SACs formed from these eight carbon supports and *3d*, *4d*, and *5d* TM elements^[80]. Using adsorption free energy of OH^{*} (ΔG_{OH^*}) as a screening criterion for ORR catalytic activity, accurately predicting the overpotentials on different carbon-based supports and establishing a volcano relationship between ΔG_{OH^*} and overpotential [Figure 4A and B]. Through systematic rapid screening, a series of carbon-based SACs with excellent ORR catalytic activity and stability were successfully identified. This study not only elucidated the influence of metal-support interactions on ORR performance but also provided strong theoretical guidance for the design and selection of carbon-supported SACs by saving experimental time and costs.

HTS can provide a valuable reference framework for screening of other multi-step reactions. Yue *et al.* investigated the nitrogen reduction reaction (NRR) performance of four surface termination structures WB₂(001) using a two-step HTS methodology [Figure 4C-F]. In the first step, they calculated the adsorption energies of N₂ and H₂, along with the free energies of the reaction pathways, to identify SACs with potential NRR activity. In the second step, they further calculated the free energies for NH₃ synthesis and desorption to select SACs with excellent NRR performance^[81]. Through this two-step screening process, they identified four notably effective SACs: V@(B-1@3), Mn@(B-2@3), Cr@(W-1@1), and Cr@(W-2@1).

Using HTS combined with first-principles calculations, researchers investigated the application of TM-tetragonal carbon nitride (TM@T-C₂N) catalysts [Figure 4G] in the electrochemical nitrate reduction reaction (NO₃RR)^[82]. They proposed a five-step screening criterion [Figure 4H], outlining the elimination sequence and stages for potential NO₃RR candidates based on different reaction phases.

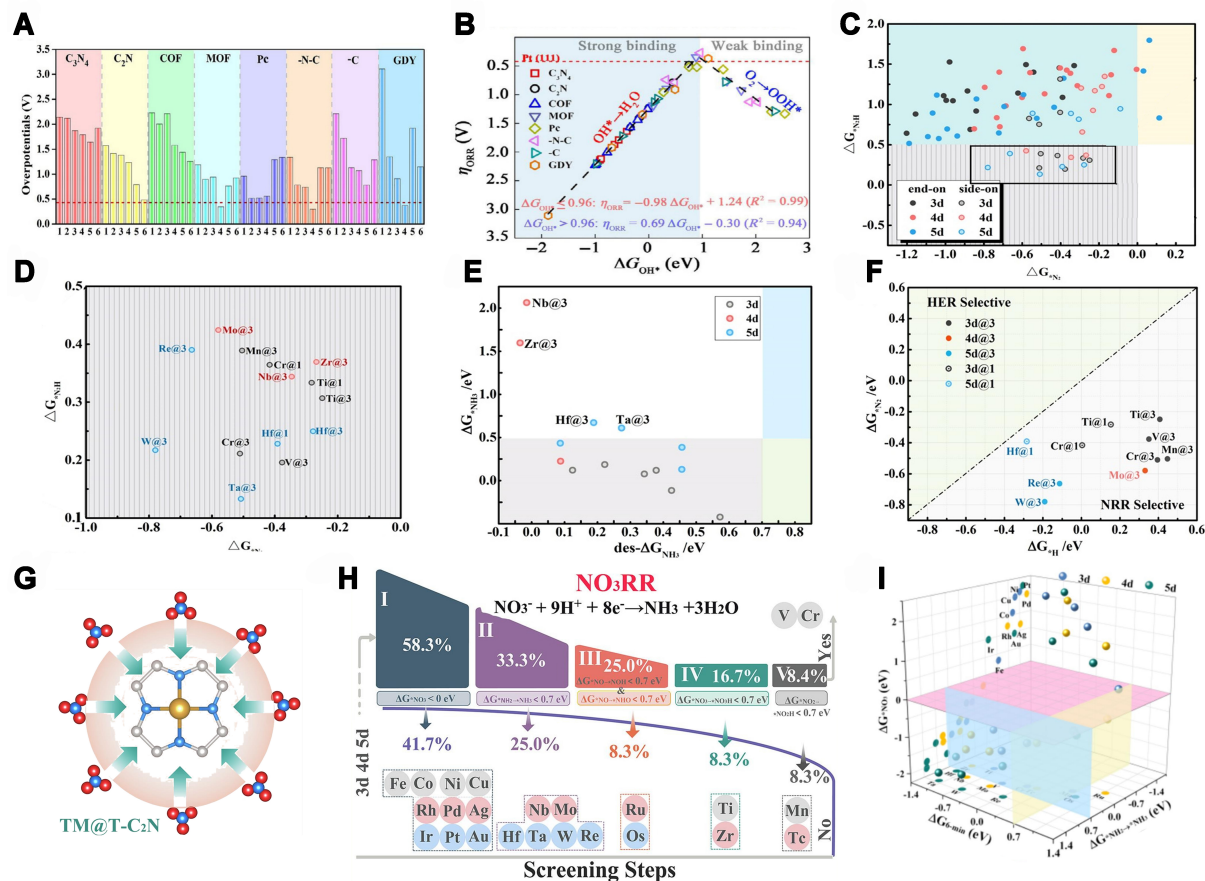


Figure 4. (A) Summary diagram of ORR overpotential for TM atom-doped different carbon-based carriers; (B) The volcano plot of ΔG_{OH^-} . The blue regions indicate strong binding energy of OH⁻; red dashed lines represent overpotential of Pt (111). Copyright 2021, Springer Nature, Reproduced with permission^[80]; (C and D) The first step of NRR screening results is to evaluate the performance of SACs supported on B-1 for N₂ fixation and conversion to N₂H; (E) The second step screening results of SAC supported on B-1 for NRR based on its NH₂-to-NH₃ formation and desorption performance; (F) Adsorption free energy diagram of H and N₂ on ten selected candidates, with white regions indicating NRR selectivity. Copyright 2022, Elsevier, Reproduced with permission^[81]; (G) The structure of TM@T-C₂N; (H) 3D screening diagram illustrating three types of free energies; (I) Schematic representation of a screening framework for identifying potential NO₃RR catalysts. Copyright 2024, Elsevier, Reproduced with permission^[82]. ORR: Oxygen reduction reaction; NRR: nitrogen reduction reaction; SACs: single-atom catalysts.

The Gibbs free energy descriptors from the first three steps were used to construct a three-dimensional (3D) screening map, as shown in Figure 4I. The pink, blue, and yellow cross-sections visualize the screening criteria: $\Delta G_{NO_3} < 0$, $\Delta G_{NH_2 \rightarrow NH_3} < 0.7$, and $\Delta G_{G-min} < 0.70$ eV, respectively. This process retained six candidate materials (Ti-, V-, Cr-, Mn-, Zr-, and Tc@T-C₂N) for further consideration. These criteria facilitated the rapid exclusion of unsuitable candidates, narrowing the focus of the study and allowing for the efficient pre-screening of a large number of candidates. Ultimately, V@T-C₂N and Cr@T-C₂N were selected as potential high-performance NO₃RR catalysts.

In a word, DFT and AIMD provide an atomic-level perspective for understanding catalyst performance, while HTS enables rapid evaluation of numerous candidates based on specific criteria, effectively reducing research time and scope. The integration of these methods offers robust support for the efficient development of SACs. It remains noteworthy that discrepancies may exist between the models employed in HTS and actual material structures, leading to deviations in the prediction results. Materials with high performance identified through screening may encounter difficulties during experimental preparation, and

the screening outcomes necessitate extensive experimental validation.

APPLICATIONS OF ML AT VARIOUS STAGES IN THE DEVELOPMENT OF SACS

Development stage: feature importance analysis using regression models

HTS generates a wealth of computational data, providing a foundational basis for subsequent analysis and model building^[83-85]. However, it does not directly assess the importance of factors influencing catalyst activity. For example, in Pt-doped Janus-MXenes, the binding energy between Pt atoms and the substrate, work function, and the number of electrons gained by Pt atoms are closely related to catalyst activity^[86]; however, the ranking of importance of different features remains unclear. Similar reports indicate that there is a strong linear relationship between the *d*-band center of metal atoms, the number of electrons transferred, and catalytic activity^[34]. However, this relationship cannot be quantified in percentage terms to express the importance of the *d*-band center.

ML, particularly regression models, can predict material performance from existing data^[87,88]. Common regression models, such as random forests, support vector machines, and decision trees, can handle multidimensional data and analyze feature contributions to catalytic performance. ML can predict new catalyst performance based on feature engineering, accelerate screening process of SACS and provide a theoretical basis for designing efficient catalysts, making research more targeted. Furthermore, the importance analysis of features deepens the understanding of how different characteristics influence catalytic performance, expediting the identification of high-efficiency catalysts.

Application of regression models to the performance analysis of SACS

Among various ML models, regression models require fewer data points and can perform feature importance analysis^[89]. Therefore, they are more suitable for integration with DFT calculations and HTS of SACS.

In this context, Chen *et al.* introduced a method for rapidly screening CO₂ reduction electrocatalysts based on simple features and ML models^[90]. The researchers constructed a database consisting of 1,060 metal-nonmetal CO-doped graphene structures, and optimized the feature set through Pearson correlation heatmaps [Figure 5A] and feature importance ranking [Figure 5B]. This led to the identification of an optimal feature set containing eight features. Various ML algorithms were tested, including K-nearest neighbors (KNN), random forest regression (RFR), support vector regression (SVR), gradient boosting regression (GBR), extreme GBR (XGBR), and a kind of composited algorithms produced by tree-based pipeline optimization tool (TPOT) [Figure 5C-H], with cross-validation used to evaluate the prediction performance of different algorithms. Among all the algorithms tested, the XGBR exhibited superior predictive performance, characterized by a higher coefficient of determination (R^2) value and a lower root mean square error (RMSE) value. When compared to the composite algorithms generated by TPOT, XGBR boasts a simpler structure, higher controllability, and greater ease in hyperparameter tuning and model optimization. The predictive model established based on the XGBR model successfully predicts changes in CO adsorption free energy (ΔG_{CO}), thereby enabling the evaluation of the catalytic activity. Moreover, the XGBR model showed excellent generalization ability. To assess the impact of the HER on CO₂ reduction, another XGBR model was developed specifically to predict the HER catalytic activity of the 1,060 materials. By combining the predictions from both models, 94 potential CO₂ reduction electrocatalysts were successfully screened.

Additionally, Zhang *et al.* employed interpretable ML models to directly predict the Gibbs free energy for evaluating the electrocatalytic nitrogen reduction activity of SACS^[91]. It is noteworthy that all the features

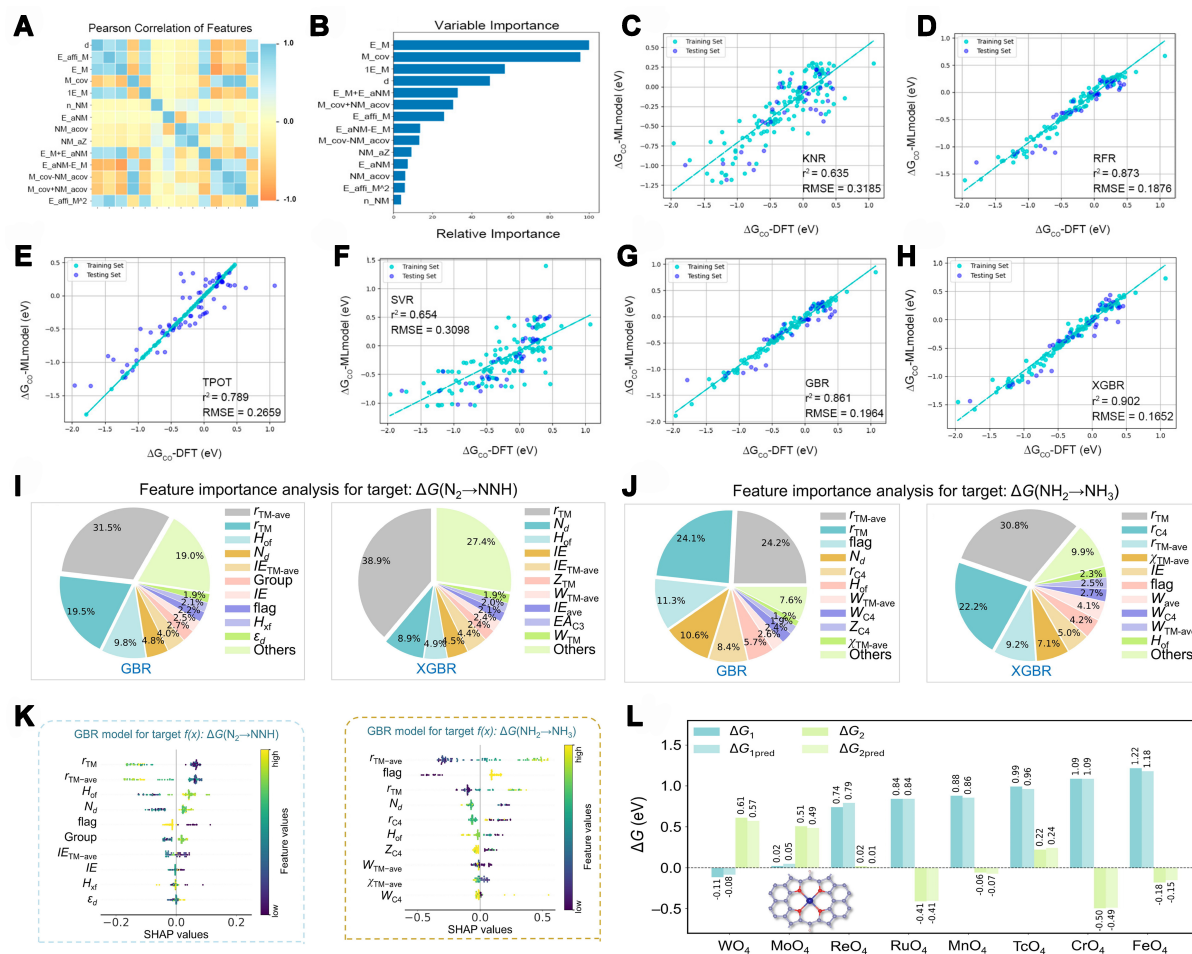


Figure 5. (A) Heatmap of Pearson correlation coefficient matrix for the ΔG_{CO} -predicted optimal feature set; (B) Ranking of feature importance within the optimal feature set; (C-H) Predictive performance of various models trained using different ML methods. Copyright 2020, American Chemical Society, Reproduced with permission^[90]; (I and J) Feature importance of the top ten significant features predicted by GBR and XGBR models; (K) Utilizing SHAP analysis to consider the overall impact of different features on model prediction; (L) Predicting reaction free energy via the GBR model: excellent agreement between predicted values and DFT calculations. Copyright 2024, Elsevier, Reproduced with permission^[91]. ML: Machine learning; GBR: gradient boosting regression; XGBR: extreme gradient boosting regression; SHAP: Shapley Additive Explanations; DFT: density functional theory.

used in the model were not derived from DFT calculations. Instead, a dataset of 90 graphene-based TM SACs was collected from available literature, which included the catalyst's structure and reaction free energies, along with 41 basic features extracted from the periodic table, such as atomic number, atomic mass, covalent radius, and *d*-electron count. Pearson correlation heatmaps were used to eliminate highly correlated features, and independent features were selected for model training. Based on the independent feature set database, various ML models, including GBR, XGBR, and RFR, were successfully trained. Ultimately, GBR was selected as the optimal ML model for predicting Gibbs free energy of NRR. It demonstrated the highest prediction accuracy during both training and testing phases with an R^2 score exceeding 0.97 and RMSE less than 0.1 eV. These performance metrics significantly outperform other models, such as XGBR and RFR. In terms of interpretability, feature importance analysis and Shapley Additive Explanations (SHAP) can uncover the working mechanisms of GBR model, providing insights into the specific influence of individual features on prediction outcomes and guiding the SAC design. Figure 5I and J show the feature importance analysis, emphasizing feature importance within the context of the model structure and training process. Figure 5K presents the SHAP analysis, which considers the overall

impact of different features on model predictions, accounting for feature interactions and correlations. GBR possesses exceptional feature capture capabilities, enabling effective identification of key features related to the active center and coordination environment, such as the radius of TM (r_{TM}), average radius of TM (r_{TM-ave}), number of d -electrons (N_d), and heat of formation (H_{of}). It was evident that the covalent radius of active center atom was consistently one of the most important features influencing NRR activity.

The analysis results based on the GBR can guide the improvement of SACs through considerations of coordination types, d -electron count, and covalent radius. For instance, selecting SACs with pyrrole-type coordination (flag = 0) can significantly reduce Gibbs free energy for NH_3 desorption, thereby enhancing NRR activity. Furthermore, by optimizing d -electron count and the difference in covalent radius, it is possible to enhance N_2 activation, suppress HER, and improve selectivity of NRR. **Figure 5L** demonstrates the predicted reaction free energies obtained through GBR align well with those calculated by DFT, validating the reliability of the model.

The studies by Chen *et al.* and Zhang *et al.* highlight the significant role of regression models in predicting catalyst performance^[90,91]. Regression models can rapidly screen potential materials, exhibiting significantly higher efficiency compared to HTS. This approach drastically reduces research costs by eliminating the need for extensive DFT calculations. Feature importance assessment can quantify the impact of different features on the catalytic performance of SACs, helping identify and focus on the most informative features. These results not only accelerate the discovery and screening process of SACs but also help understand the internal mechanisms of ML models. Furthermore, the trained ML models exhibit excellent transferability and robustness, making them powerful tools for future catalyst research and development.

Feature importance analysis accelerates the screening process of SACs and provides a theoretical foundation for designing efficient catalysts, making the research process more targeted^[92-94]. In a recent study, Pritom *et al.* combined DFT and ML methods to investigate the influence of porphyrin-supported SACs on the charging behavior of non-aqueous magnesium-carbon dioxide (Mg-CO₂) batteries^[95]. The study reveals the high efficiency of SACs in both the binding of MgCO₃ and promoting its decomposition. By employing a ML model, the adsorption energy of MgCO₃ on SACs was successfully predicted, and the key factors influencing the adsorption energy were analyzed. The results indicate that the ionization potential of the TM is a crucial parameter for selecting SACs.

In the process of constructing the ML model, the authors selected electronic structure features closely related to the adsorption performance of SACs as inputs. These features include the distance between the TM and N/S, the charge of the TM, the N_d , the average charge of nitrogen, electronegativity, ionization energy, and electron affinity. Correlation analysis using the Pearson correlation coefficient was performed to select independent features and identify redundant ones. For example, it was found that the TM charge was significantly correlated with other features, so it was removed from the model. Additionally, the authors used violin plots [**Figure 6A** and **B**] to show how the adsorption energy of MgCO₃ exhibited different distributions at various levels under different environments, clearly revealing the heterogeneity of the data distribution and the differences across configurations, providing an intuitive basis for feature selection.

To predict the adsorption energy, the authors selected two models, GBR and artificial neural networks (ANNs), due to their advantages in handling nonlinear relationships and complex data. Prior to model training, the data was standardized to ensure comparability between different features, and K-fold cross-validation was employed to assess the model's generalization ability and avoid overfitting. By training the GBR and ANN models with DFT-calculated results and adjusting the model parameters and structures, the

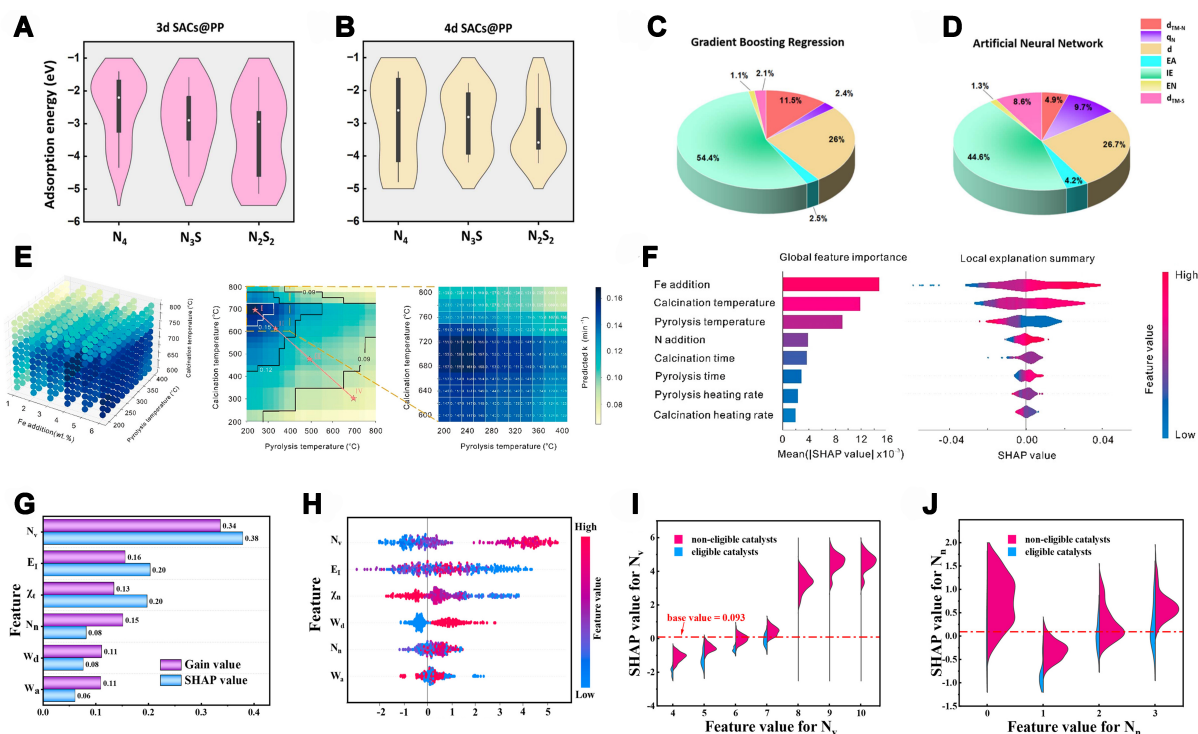


Figure 6. (A and B) Violin plots illustrating the adsorption energies of MgCO_3 across various environments; (C and D) Graphic representation of feature importance utilizing MDI from GBR model and permutation importance technique from ANN model. Copyright 2024, Royal Society of Chemistry, Reproduced with permission^[95]; (E) Representation of predictions generated by a training model using a 3D matrix composed of the three most influential descriptors, with four distinct regions classified by k values, highlighting a matrix of calcination and pyrolysis temperatures and enlarging the area with high k values; (F) Violin plot of SHAP values for the XGBR model (left) and summary plot of SHAP analysis (right). Copyright 2023, American Chemical Society, Reproduced with permission^[41]; (G) Illustrating the cruciality of six features through gain value and SHAP value; (H) SHAP summary plot for ML models; (I and J) Violin plots of SHAP values for N_v and SHAP dependence for N_n , with blue indicating eligible catalysts and red indicating ineligible ones. Copyright 2023, John Wiley and Sons, Reproduced with permission^[100]. MDI: Mean decrease in impurity; GBR: gradient boosting regression; ANN: artificial neural network; SHAP: Shapley Additive Explanations; XGBR: extreme gradient boosting regression; ML: machine learning.

authors achieved high prediction accuracy.

To gain deeper insights into model predictions, the authors employed multiple methods such as SHAP, permutation importance, and mean decrease in impurity (MDI) [Figure 6C and D] to analyze feature importance. These methods evaluated the influence of each feature on the model's predictions from different perspectives and revealed that ionization energy and the N_d were key features influencing the MgCO_3 adsorption energy. It suggests that future improvements in the catalytic performance of SACs can be achieved by optimizing these key features.

For the optimization of multi-step chemical transformations, a ML framework has been developed to guide catalyst design by analyzing key steps in the multi-step process to enhance reaction efficiency^[96-99]. In another study, Fu *et al.* reported the use of ML algorithms to accelerate the design process of highly efficient Fenton-like SACs^[41]. The XGBR prediction model built using ML algorithms accurately predicted the degradation rate (k -value) of SACs for phenol. The SHAP explanation method quantified the impact of various parameters on the model's predictions, as shown in Figure 6E and F, revealing that Fe loading, carbonization temperature, and carbonization heating rate are key factors influencing the k -value. Through ML-guided optimization, they identified efficient SACs dominated by Fe- N_s sites, exhibiting excellent

Fenton-like activity ($k = 0.158 \text{ min}^{-1}$). This work provides an example of ML-assisted optimization of single-atom coordination environments and demonstrates its feasibility in accelerating the development of high-performance catalysts, thereby helping researchers gain a deeper understanding of the structure-performance relationship of SACs.

Materials with high performance identified through screening may encounter difficulties during experimental preparation, and the screening outcomes necessitate extensive experimental validation. Sun *et al.* proposed a strategy based on interpretable ML and HTS to understand and predict the performance of NRR catalysts^[100]. Through DFT calculations, a catalyst space containing 168 types of carbon-supported SACs was constructed. A four-step screening strategy was employed to identify 33 promising candidate catalysts. Using the interpretable ML model XGBR, the key features influencing NRR performance were analyzed. The feature importance, SHAP value distribution, and dependency plots [Figure 6G-J] revealed that the valence electron count (N_V) and nitrogen substitution number (N_n) of the metal atom are the two key factors affecting catalytic performance. Further analysis showed that active centers with lower N_V and moderate N_n could better balance charge distribution, thereby enhancing NRR catalytic activity. Ultimately, six catalysts with excellent NRR performance were selected, and reaction pathways and rate-determining steps were analyzed.

To sum up, ML models are crucial for SAC studies, enabling precise prediction and optimization of catalyst performance, thus enhancing design efficiency and understanding of catalytic mechanisms. AI methods rely heavily on data quality and quantity, which are costly and hard to obtain. Many ML models, especially deep learning, lack interpretability, making it difficult to understand the physical mechanisms behind predictions. Additionally, their generalization ability may be limited when faced with new, unseen data. As ML advances, continuous optimization of algorithms and models will improve accuracy and reliability, driving the progress of catalytic science.

Growth stage: using NN to analyze the structural characteristics of SACs and screen high-performance catalysts

By using ML regression models, researchers have successfully identified key features that influence catalyst activity and conducted feature importance analysis^[36,101-103]. A critical challenge now is leveraging these important features to screen and predict the catalytic performance of new structures, which is key to simplifying the screening process and reducing costs.

NNs [such as deep neural networks (DNNs), convolutional neural networks (CNNs), and graph neural networks (GNNs)] are particularly well-suited for this task due to their layered structure, with each layer containing multiple neurons. These networks can process input data through activation functions, capturing complex nonlinear relationships. By automatically learning high-dimensional features, NNs can better describe and predict the properties of materials^[104], making them ideal for handling complex models and large-scale datasets. Consequently, NNs hold great promise for screening high-performance catalysts from complex structures.

DNNs can significantly reduce computational time, quickly eliminating large numbers of ineffective catalysts, and saving computational resources^[105-107]. At the same time, these models help reveal the relationship between structural features and reaction mechanisms^[39,108]. Zafari *et al.* utilized coulomb matrices and principal component analysis (PCA) to reduce the dimensionality of geometric structure features of SACs, which helped extract relevant features and reduce model complexity^[109]. The optimized structural information, consisting of seven features, was used as the input for the DNN model with an input layer, two hidden layers, and an output layer [Figure 7A]. The DNN model predicts N_2 adsorption energy

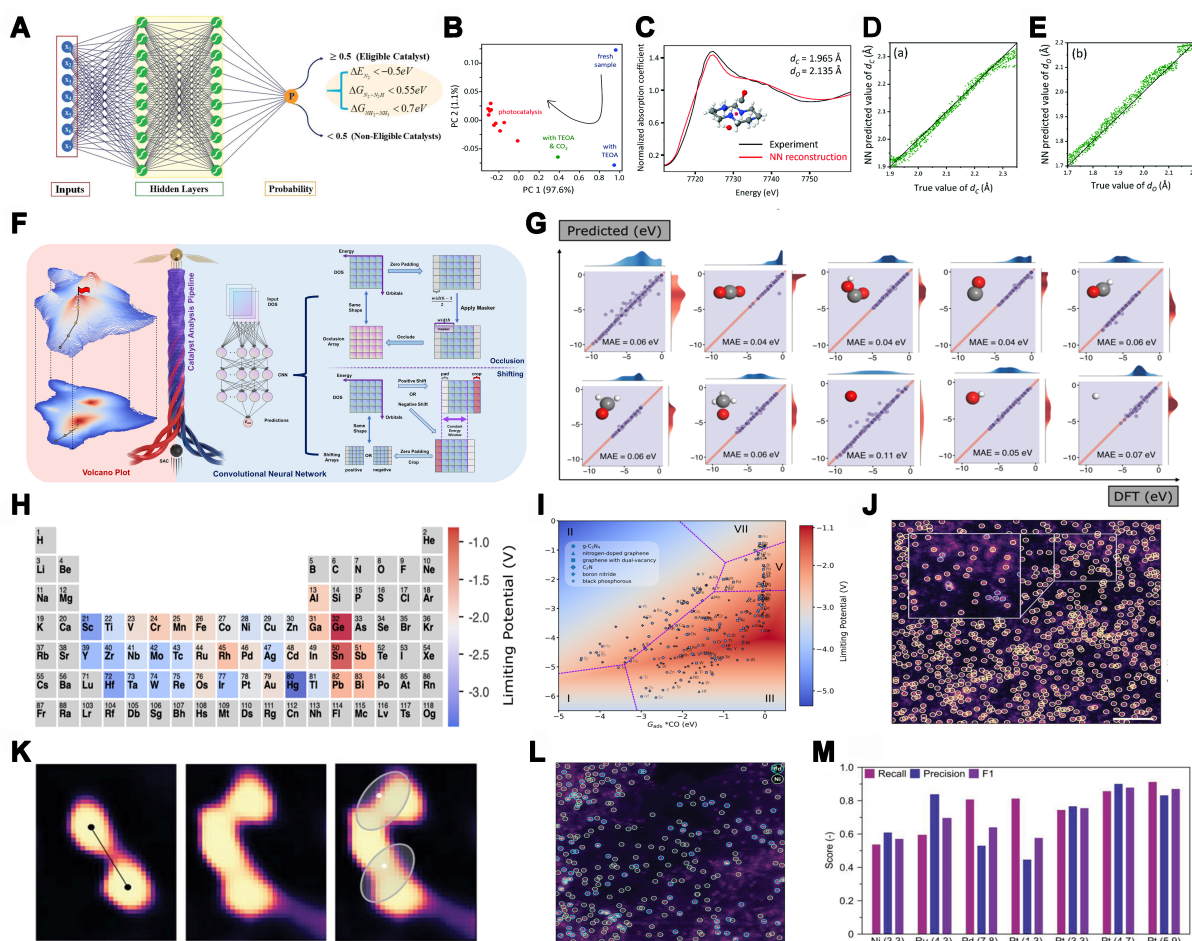


Figure 7. (A) An illustrative diagram of the ANN architecture (featuring 10 neurons per hidden layer), utilizing optimized SAC geometric models as input data, with each structural geometry possessing seven distinct features. Copyright 2020, Royal Society of Chemistry, Reproduced with permission^[109]; (B) Classification of SACs into three categories using PCA and K-means clustering of XANES data (with different colors representing distinct clusters); (C) Comparison of experimental XANES spectra (in black) with theoretical XANES spectra reconstructed from descriptor values predicted by NN; (D and E) The prediction accuracy of NN models on the test dataset. Copyright 2022, Royal Society of Chemistry, Reproduced with permission^[110]; (F) Volcano plot and CNN-based catalyst performance analysis pipeline: integrative use of volcano plots for predictive assessment of existing catalysts, with eDOS as input for predicting and tuning adsorption energies, and extraction of chemical information from CNN model; (G) Prediction of adsorption energies for intermediates in the CO₂RR process using a CNN model; (H and I) Limiting potential volcano plot and periodic table. Copyright 2024. This publication is licensed under CC-BY-SA 4.0^[112]; (J) AC-STEM image of Pt₁/NC with ML-detected overlapping atoms highlighted in yellow circles. Magnified view emphasizing potential elements for detection and quantification (indicated by blue circles); (K) The representative prediction maps generated by CNN for elements in (J), with interatomic distance analysis in the left image and overlapping features addressed using Gaussian Mixture Models assignment in the two images on the right; (L) Inference of corresponding atomic position assignments by CNN models and prediction of atomic chemical properties through VAE latent space clustering; (M) Comparison of performance metrics for ML models versus manual tasks executed by domain experts: ML model detection on test images accomplished in minutes versus hours required for human expert tasks. Copyright 2023, John Wiley and Sons, Reproduced with permission^[116]. ANN: Artificial neural network; SAC: single-atom catalyst; PCA: principal component analysis; XANES: X-ray absorption near-edge structure; NN: neural network; CNN: convolutional neural network; eDOS: electronic density of states; AC-STEM: aberration-corrected scanning transmission electron microscopy; ML: machine learning; VAE: variational autoencoders.

and hydrogenation free energy, successfully screening three NRR electrocatalysts-CrB₃C₃, TcB₃C₃, and HfB₃C₂ with low overpotential and high selectivity.

X-ray absorption near-edge structure (XANES) analysis is a powerful technique for probing the structural changes in SACs. However, traditional XANES analysis struggles to handle structural heterogeneity, making it challenging to accurately identify the number of SAC species. In the study by Xiang *et al.* on photocatalytic CO₂ reduction, the NN model could extract structural information from XANES data, enabling a more accurate and efficient identification of SAC species and structural variations during reaction processes^[110].

Using methods such as PCA, K-means clustering, and NNs, the XANES data was successfully analyzed to obtain quantitative structural information about the local atomic environment of SACs. This approach allowed the classification of SAC species into three categories [Figure 7B], identifying their number and providing a foundation for subsequent structural analysis. By employing the NN-XANES method, the local geometry of Co-cyclaml-CO was refined. A large set of theoretical XANES data was trained to establish a mapping relationship between the XANES features and structural descriptors. The trained NN model showed consistency with experimental data [Figure 7C-E], and it could predict structural descriptors from experimental XANES data, such as bond lengths and bond angles. Additionally, it provided more detailed structural information, allowing for the distinction of Co-O and Co-N contributions.

CNNs are a type of feedforward NN composed of convolutional layers, pooling layers, and fully connected layers^[111]. With features such as local connections, weight sharing, and pooling, CNNs excel in processing images or structural data of materials. Unlike traditional methods such as the *d*-band center model, CNNs automatically extract features from electronic DOS (eDOS) without manual intervention, establishing complex relationships with adsorption energy. Yang *et al.* proposed a workflow that combines CNNs with volcano plots [Figure 7F] to screen and predict two-dimensional SACs in CO₂RR^[112]. By establishing correlation plots between intermediate adsorption energies and various descriptors, a CNN model was used, with 2D eDOS as input, to predict adsorption energies and understand the impact of electronic structure perturbations. The CNN model predicted the adsorption energies of nine intermediates in CO₂RR with an average mean absolute error (MAE) of 0.06 eV [Figure 7G], demonstrating high prediction accuracy compared to DFT. It also exhibited strong generalization ability in handling species containing oxygen, hydrogen, carbon atoms, and different substrates. In addition, a hybrid descriptor combining C-type and O-type CO₂RR intermediates was introduced to construct optimized volcano plots and periodic tables [Figure 7H and I], providing a visual approach to efficiently screen for CH₄ reduction catalysts.

CNNs also play a crucial role in the structural characterization of SACs, enabling rapid, accurate, and automated detection of metal centers^[113-115]. Due to the current lack of research focused on metal centers, it is challenging to design atomically precise structural materials. However, the use of CNNs enables rapid and standardized detection of metal centers^[116]. CNNs can identify pixel patterns in aberration-corrected scanning transmission electron microscopy (AC-STEM) images, distinguishing between metal atoms and background pixels accurately. Threshold segmentation and bounding box recognition techniques allow for thresholding of the probability map output by CNNs and the use of bounding boxes to identify the coordinates of metal atoms [Figure 7J and K]. CNNs and Gaussian mixture models can perform chemical specificity analysis on multi-metal ultra-high-density (UHD)-SACs [Figure 7L], distinguishing metal centers of different chemical types and quantifying mixing degrees between metal centers. Compared to manual methods, CNNs demonstrate higher accuracy and repeatability, as shown in Figure 7M, greatly improving detection efficiency.

GNNs further expand the application of NNs and are tailored for processing graph-structured data, particularly useful for handling molecular or crystal structures^[117]. By representing material structure as

graph networks with nodes (atoms) and edges (bonds), GNNs can efficiently extract both local and global structural information, propagating information between adjacent nodes and learning complex structural relationships^[118]. For SACs, GNNs can model the interactions between metal atoms and supports, predicting catalytic performance based on these interactions.

In material design, surface structural changes at the nanoscale are especially important. Small molecule adsorption energy is a key indicator of catalyst activity, but linear scaling relations limit performance improvement. Surface strain can break these scaling relationships. Surface strain engineering involves a high-dimensional search space, and comprehensive DFT screening is impractical. GNNs can efficiently handle high-dimensional data by learning nonlinear functions and generalizing from relatively small training datasets, allowing for efficient exploration of the strain space. Using GNNs to predict the adsorption energy response of catalyst/adsorbate systems under surface strain patterns, Price *et al.* proposed a GNN model [Figure 8A], effectively bridging the gap between experimental and theoretical results^[119]. The normalized confusion matrix of the GNN + strain model for experimental data [Figure 8B] successfully predicts the strain patterns in 85% of unseen test data, outperforming linear models. The model also predicted strain responses in ammonia synthesis reaction intermediates [Figure 8C], revealing the role of compressive strain in breaking linear scaling relations. This study provides a new approach for identifying strain patterns that can break the adsorption energy scaling relationships. By generating phase diagrams of adsorption energy versus strain [Figure 8D and E], it offers an intuitive method for strain engineering, guiding catalyst design and improving performance.

It is clear that both CNN and GNN models have their own strengths. The crystal graph convolutional neural network (CGCNN) combines the advantages of both CNNs and GNNs, enabling it to learn material properties from atomic connections within crystals and providing highly accurate predictions. CNNs excel at processing image data and can be used to identify the crystal structures of materials. In contrast, GNNs are well-suited for handling graph-structured data to capture atomic interactions and chemical bonding information.

Figure 8F indicates that the CGCNN accelerates high-performance dual-atom catalyst (DACs) screening by learning the structure-activity relationships of existing DACs, enabling the prediction of ΔG_{H^*} and facilitating the selection of graphene-based DACs with ideal conductivity and stability for HER, thereby reducing the computational cost of DFT by half^[120]. By combining CGCNN with DFT [Figure 8G], the model screened 435 dual-atom combinations of nitrogen-doped graphene (N_6Gr) and identified two high-performance HER catalysts ($AuCo@N_6Gr$ and $NiNi@N_6Gr$). DFT confirmed these two catalysts exhibited excellent reaction kinetics, and an analysis of electronic states, ΔG_{H^*} , and metal-hydrogen interaction strength revealed the crucial role of metal synergy in HER.

CGCNN avoids the need for feature engineering by directly learning the chemical structure from the material's geometric configuration, which simplifies the screening process and saves computational resources. The ΔG_{H^*} predicted by the CGCNN model shows a high agreement with the DFT results, with a MAE and RMSE of 0.22 and 0.27 eV, respectively [Figure 8H], demonstrating the model's high accuracy. The CGCNN is effectively applied to the screening of DACs, contributing to the acceleration of catalyst discovery.

Overall, both CNNs and GNNs offer significant advantages in analyzing complex catalyst structures. By learning high-dimensional features automatically, they improve the description and prediction of material properties, advancing materials science. These methods help researchers identify promising high-

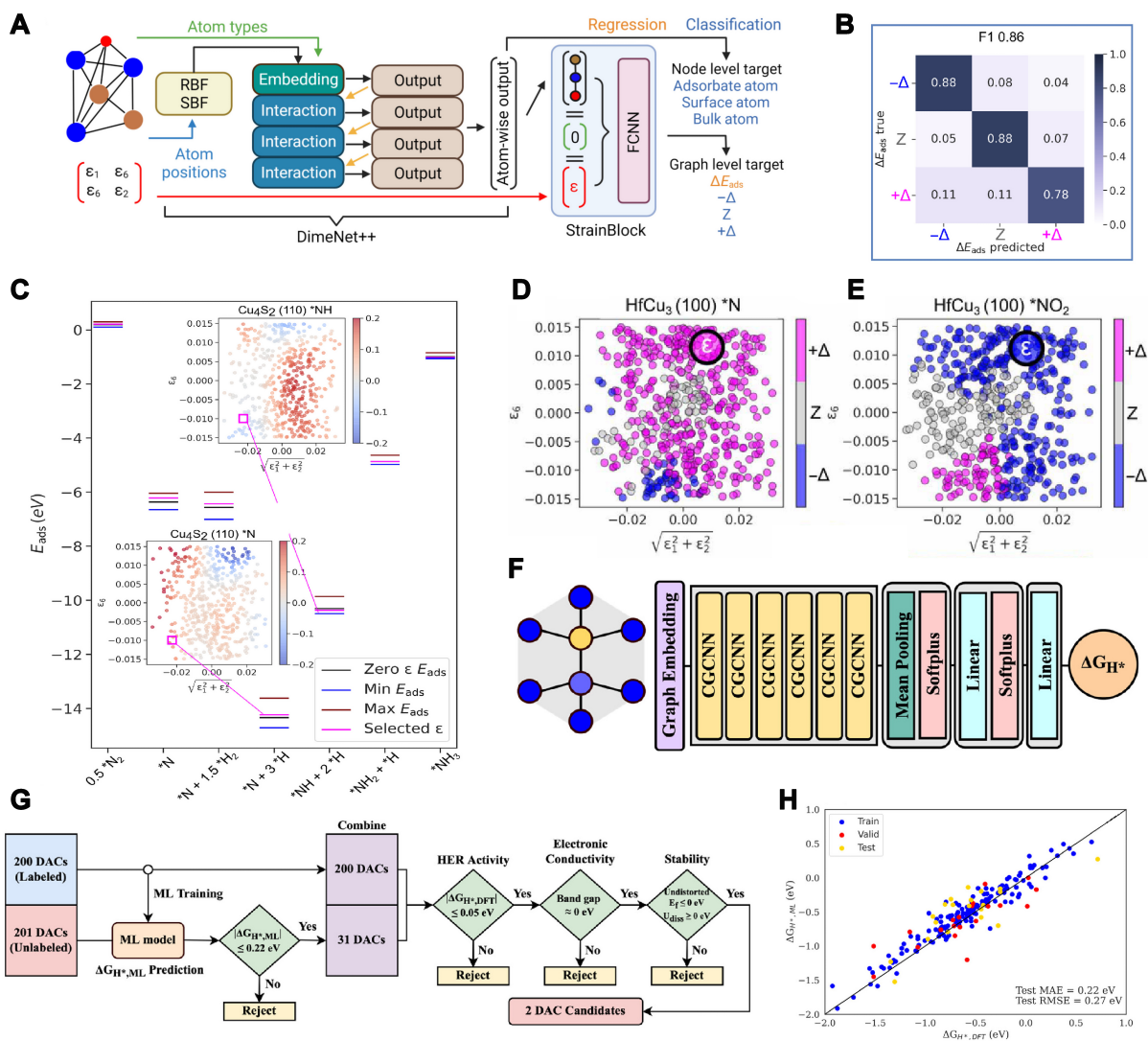


Figure 8. (A) The architecture of GNNs for classification and regression tasks; (B) The normalized confusion matrix for test data. Each row corresponds to different true classes, and each column corresponds to predicted classes; the diagonal represents the percentage of correct predictions for each class; (C) Prediction diagram of strain response for single-molecule NH_3 synthesis on $\text{Cu}_4\text{S}_2(110)$ surface by regressor; (D and E) Verification and comparison of strain phase diagrams for $\text{HfCu}_3(100)$ surface adsorption with N and NO_2 using DFT. Copyright 2022, The American Association for the Advancement of Science, Reproduced with permission^[119]; (F) Diagram of the GNN architecture applied in this study; (G) The workflow diagram of combining ML with DFT screening; (H) The pairing diagram of CGCNN model, DFT-calculated ΔG_{H^*} , and ML-predicted ΔG_{H^*} . Copyright 2023, American Chemical Society, Reproduced with permission^[120]. GNNs: Graph neural networks; DFT: density functional theory; ML: machine learning; CGCNN: crystal graph convolutional neural network.

performance catalysts from numerous SAC combinations, significantly reducing experimental efforts. Future work should address the high computational costs and complex training processes of NNs for large-scale material structures. Additionally, attention is needed to prevent over-smoothing in deep, multi-layer networks, as it can reduce model performance.

Maturity stage: design the SAC structure and predict its performance with a generation model

In the mature phase of integration between AI and materials science, generative network models have demonstrated immense potential in catalytic structure design. In this reverse design process, users can define catalytic properties of SACs and generate model structures with precisely defined attributes. For

instance, generative models such as GANs and variational autoencoders (VAE) can learn low-dimensional representations from training data and continuously alter parameters^[121,122]. These models coupled with computational methods such as DFT to validate whether the generated structures meet required performance criteria, thereby optimizing the design. In specific applications, generative models extract synthesis steps and catalytic properties from literature, employ active learning to explore the chemical space of specific catalysts, and leverage models such as GANs and VAE to generate hypothetical alloys and ligands^[123].

Generative models offer an efficient, diverse, and interpretable approach to swiftly generating and evaluating a variety of catalyst structures, thus accelerating the discovery and design process^[124,125]. GANs can help automate the improvement of catalyst materials by generating new catalyst surfaces with higher activity. Ishikawa proposed a novel approach that combines computational chemistry and ML to “extrapolate” new catalytic surfaces^[126]. DFT is used to calculate the energy of basic reactions on a given set of catalyst materials, and then the results are input into the GANs.

Using a GAN trained on a DFT dataset, researchers have successfully generated more complex and diverse catalyst structures, expanding the possibilities for catalyst design [Figure 9A]. These newly generated surface structures, not included in the initial dataset, exhibit higher turnover frequency (TOF) values for the ammonia synthesis reaction. Through iterative training [Figure 9B], the model continuously learns patterns and trends from the DFT dataset, applying them to create novel catalyst surface structures. This approach leverages GANs in combination with DFT calculations to “extrapolate” catalysts with enhanced catalytic activity, optimizing key factors such as reaction energy and activation energy.

In this method, DFT calculations are used to compute the energy (ΔE) of elementary reactions for all surfaces present in the initial dataset. TOF values for ammonia synthesis are then obtained from ΔE values, and metal surfaces are labeled according to their TOF. A GAN, composed of a discriminator and generator, is trained on this DFT dataset, enriched with TOF values and metal surface information.

The generator of the GAN creates samples not contained in the current dataset. In this case, a conditional GAN is employed to generate surfaces with higher TOF values. Figure 9C demonstrates the GAN to learn and recognize key factors affecting catalyst activity, such as step sites adjacent to adsorbing atoms and to apply this knowledge to produce new catalyst surface structures with enhanced activity. DFT calculations are then performed on the newly generated samples, with results integrated back into the dataset. Figure 9D and E illustrate the iterative process, starting with 100 random steps and alloy surfaces created through atomic substitution. After five iterations, a previously unobserved $\text{Rh}_8\text{Ru}_{76}$ surface was successfully obtained, achieving a TOF more than ten times higher than the best TOF value in the original dataset.

All in all, samples generated in later iterations typically exhibit higher TOF values, indicating that the iterative DFT-GAN approach effectively enhances NN training within the GAN. Moreover, the generated surfaces tend to show a higher proportion of Ru atoms, aligning with experimental observations. This improvement is attributed to a lower activation energy for the RDS due to the reduced dissociation energy of N_2 and a lower formation energy of NH_3 , which decreases NH_2 coverage on the surface and mitigates NH_2 poisoning. These characteristics contribute to higher TOF values in the generated surfaces. This study demonstrates that combining DFT with GAN is a promising strategy for the automated, continuous improvement of catalyst performance. Compared to traditional catalyst design methods, GANs enable the rapid and efficient generation of novel catalyst surfaces with higher catalytic activity, facilitating catalyst material optimization without the need for manual intervention, thus enhancing both the efficiency and

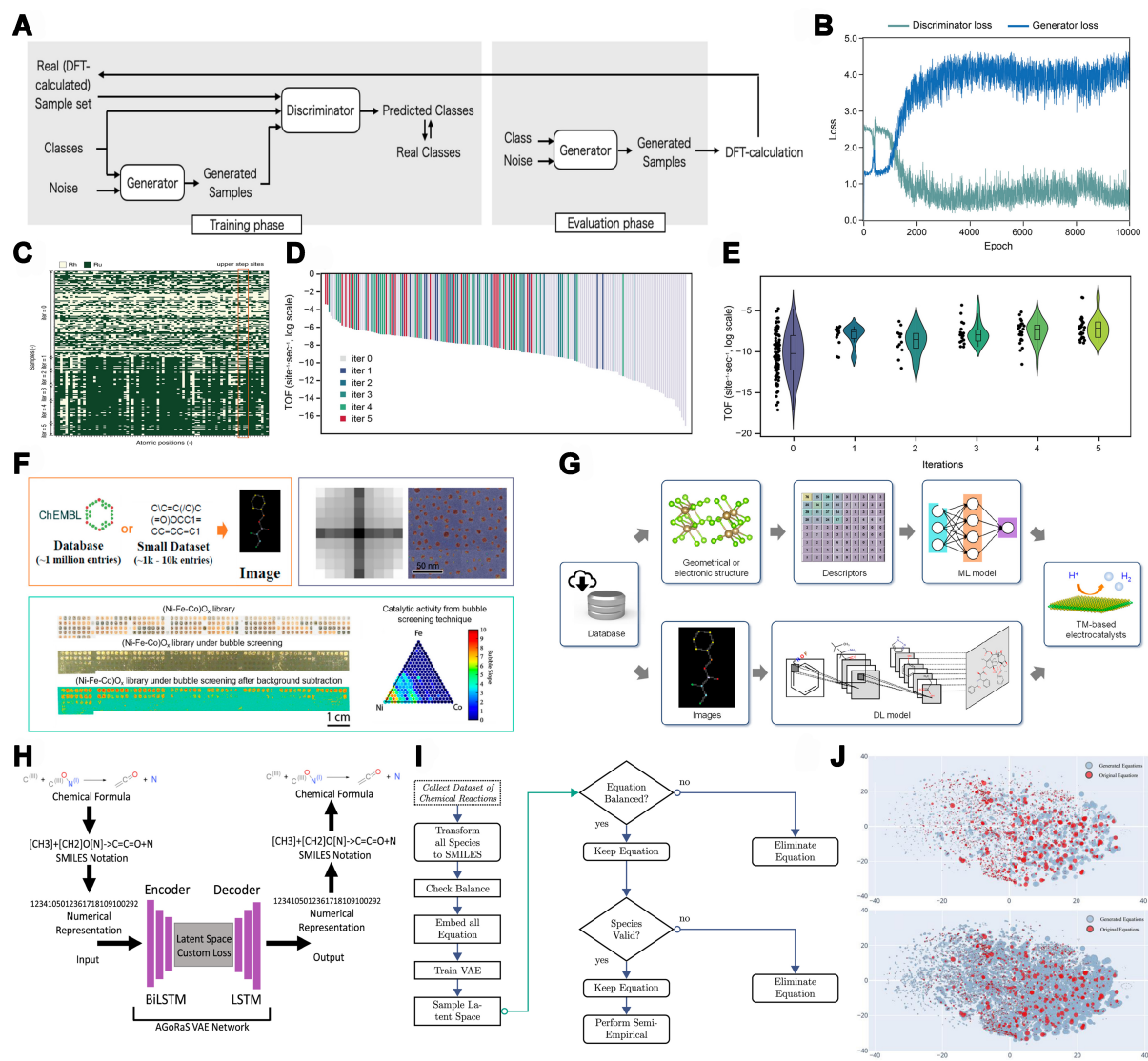


Figure 9. (A) Flow chart of the DFT-GAN program structure: training and evaluation phases; (B) Discriminator and generator losses in DFT-GAN, with each iteration comprising 2,000 epochs; (C) The distribution maps of Ru and Rh atoms on the initial surface (iter = 0) and GAN-generated surfaces (iter = 1-5); (D) The TOF of NH₃ formation on Rh-Ru alloy surface, with TOF values from different DFT-GAN iterations (iter = 1-5) encoded in distinct colors; (E) Box plot and violin plot of TOF values for iter = 0-5. Left-side points on violin represent raw TOF values for each iteration. Copyright 2022, Springer Nature, Reproduced with permission^[126]; (F) Application potential of DL techniques in image-based catalyst screening, with recognizable image types including chemical images, morphological images, and catalytic images; (G) The workflow diagram of ML and DL for discovering HER electrocatalysts. Copyright 2021, American Chemical Society, Reproduced with permission^[127]; (H) Flow chart of the VAE network in AGoRaS: decomposition of chemical database information into a high-dimensional latent space; (I) The flow chart of data collection, training, and validation steps employed by AGoRaS for generating synthetic data; (J) t-SNE visualizations of training and generated datasets. Upper panel: a sample of 7,000 equations from the training dataset, alongside 7,000 randomly selected equations from the generated dataset. Lower panel: representation of 70,000 equations extracted from the generated dataset. Copyright 2022, Springer Nature, Reproduced with permission^[128]. DFT: Density functional theory; GAN: generative adversarial network; TOF: turnover frequency; DL: deep learning; ML: machine learning; HER: hydrogen evolution reaction; VAE: variational autoencoders; t-SNE: t-distributed stochastic neighbor embedding.

accuracy of catalyst design.

Deep learning image-based recognition offers distinct advantages in processing various types of image data, such as chemical, morphological, and catalytic images. [Figure 9F](#) highlights the potential of deep learning

techniques in image-based catalyst screening. These images can provide valuable information, enabling researchers to rapidly identify and screen efficient TM-based electrocatalysts for HER. By introducing the latest advancements for identifying highly active TM-based HER electrocatalysts, representative studies utilize deep learning NN architectures [Figure 9G] to screen catalysts based on chemical and morphological images. This approach enhances the understanding of the relationship between the intrinsic properties of TM-based materials and their electrocatalytic performance^[127].

This generative design approach transforms AI from merely a tool for predicting material properties to an active driver in the discovery of new materials. Tempke and Musho have introduced an AI model known as AGoRaS, based on a VAE, designed for synthesizing new chemical reaction datasets^[128]. The AGoRaS VAE model is structured with embedding layers, bidirectional long short-term memory (LSTM), and latent space sampling, enabling the generation of an unbiased chemical reaction dataset by encoding reaction data into a latent space [Figure 9H]. By sampling within this latent space, the model generates novel chemical reactions, circumventing the biases typically present in traditional datasets. The model incorporates a rigorous data collection, training, and validation pipeline to ensure the validity and stability of generated reactions [Figure 9I].

Figure 9J illustrates a t-distributed stochastic neighbor embedding (t-SNE) plot that reduces high-dimensional data to two dimensions, showcasing the model's application on a training dataset of 7,000 reactions. This approach successfully produced over seven million new reactions, including 20,000 molecular species previously absent from the dataset, broadening the predictive capabilities of ML algorithms. Additionally, the AGoRaS model predicts thermodynamic properties for new reactions, such as Gibbs free energy, entropy, and dipole moments. Semi-empirical calculations help validate the stability of the predictions. By enabling targeted reaction searches that include specific molecular species, this model provides new avenues for experimental research. AGoRaS is thus poised to facilitate the generation of novel chemical reactions, guiding experimental efforts while enhancing the robustness and accuracy of AI algorithms in chemical science.

ML, coupled with genetic algorithms and Bayesian optimization, enables the continuous evolution of catalyst structures, optimizing performance. This approach accelerates SAC discovery and facilitates custom designs for clean energy and environmental protection. By leveraging structure-performance databases, GANs generate efficient catalysts with novel atomic arrangements, enhancing activity and stability for specific reactions such as CO₂ reduction and oxygen reduction.

The demand for novel functional materials necessitates effective strategies to expedite material discovery, where crystal structure prediction is foundational^[129,130]. Generative models such as GANs serve as a powerful means to explore hidden regions within chemical space^[122]. Kim *et al.* proposed a GAN-based approach for crystal structure prediction^[131], structured with a generator, discriminator, and classifier. The generator produces new crystal structures from random noise vectors and encoded composition vectors; the discriminator computes the Wasserstein distance between real and generated data to assess authenticity, and the classifier ensures the generated structures align with specified composition. Applying the GAN model to Mg-Mn-O systems enabled generative high-throughput virtual screening for photoanode properties. The crystal structures generated by the GAN model demonstrated reasonable stability and Heyd, Scuseria, and Ernzerhof (HSE)-calculated band gaps, some of which possess unique configurations compared to existing materials.

Another innovative approach, the CGCNN, has been developed for predicting the properties of crystalline materials^[132]. This method represents crystal structures as graphs and builds CNN upon them [Figure 10A]. CGCNN learns and automatically extracts atomic connectivity features within crystal structures to predict various material properties, such as formation energy, band gap, and elastic properties. By leveraging CGCNN, researchers can estimate energy contributions of atoms within perovskites, uncovering empirical design rules of materials.

In summary, generative models offer distinct advantages in exploring chemical space, especially in uncovering regions that traditional methods miss. They enable researchers to evaluate the impact of local chemical environments on global properties and perform combinatorial searches for synthesizable materials, providing valuable insights and design rules. This approach narrows the search and accelerates discovery.

GANs can pinpoint local maxima within the material design space, guiding materials toward optimal performance. For SACs, catalytic properties are closely linked to the local environment of metal atoms on support materials, including electronic structure, atomic spacing, and coordination numbers^[133,134]. By simulating these local environments, GANs can generate catalyst models with ideal electronic structures, achieving improved activity and selectivity in catalytic processes. Moreover, the generative and discriminative mechanisms of GANs allow the swift screening of numerous candidate structures, making this approach especially effective for optimizing heterogeneous catalysts, where traditional experimental screening would be prohibitively costly and time-intensive.

On the data-driven front, AI/ML can develop performance surrogate models that generate catalyst structures with specific chemical properties or reaction pathways based on given input conditions^[135]. These models can also predict the properties of unknown substances by training NNs to correlate input parameters with catalytic performance, enabling the identification of new active sites and guiding the optimization of experimental conditions [Figure 10B]. However, the field faces challenges, such as system complexity, data diversity, and target variability. To address these, techniques such as genetic algorithms [Figure 10C] and Bayesian optimization are explored to guide optimization, providing structured approaches to refine catalyst design amidst these complexities^[136].

NNs face limitations, including extensive data and computational needs, and lack of interpretability, necessitating integration with other technologies. High-entropy alloy catalysts are notable for superior catalytic performance; however, traditional trial-and-error methods hinder systematic exploration of structure-performance relationships. By training NNs on small unit structures and using inverse design for larger structures [Figure 10D], high-entropy alloy models can be efficiently generated to analyze structure-performance relationships. NNs predict key catalytic features, such as adsorption energies at active sites, identifying the most influential sites. Combined with Bayesian optimization [Figure 10E], this accelerates discovery of high-performance high-entropy alloy catalysts by automatically identifying optimal compositions and predicting activities^[137].

An innovative AI-driven catalyst optimization workflow, incorporating large language models (LLMs), Bayesian optimization, and active learning loops, can accelerate catalyst optimization^[138-140]. This workflow effectively combines advanced language understanding with robust optimization strategies, translating knowledge extracted from diverse literature into practical parameters for experimentation and optimization. Lai *et al.* demonstrated the application of AI workflow in the synthesis of ammonia synthesis catalysts; the results indicated that this workflow simplifies the catalyst development process^[141]. Figure 10F and G

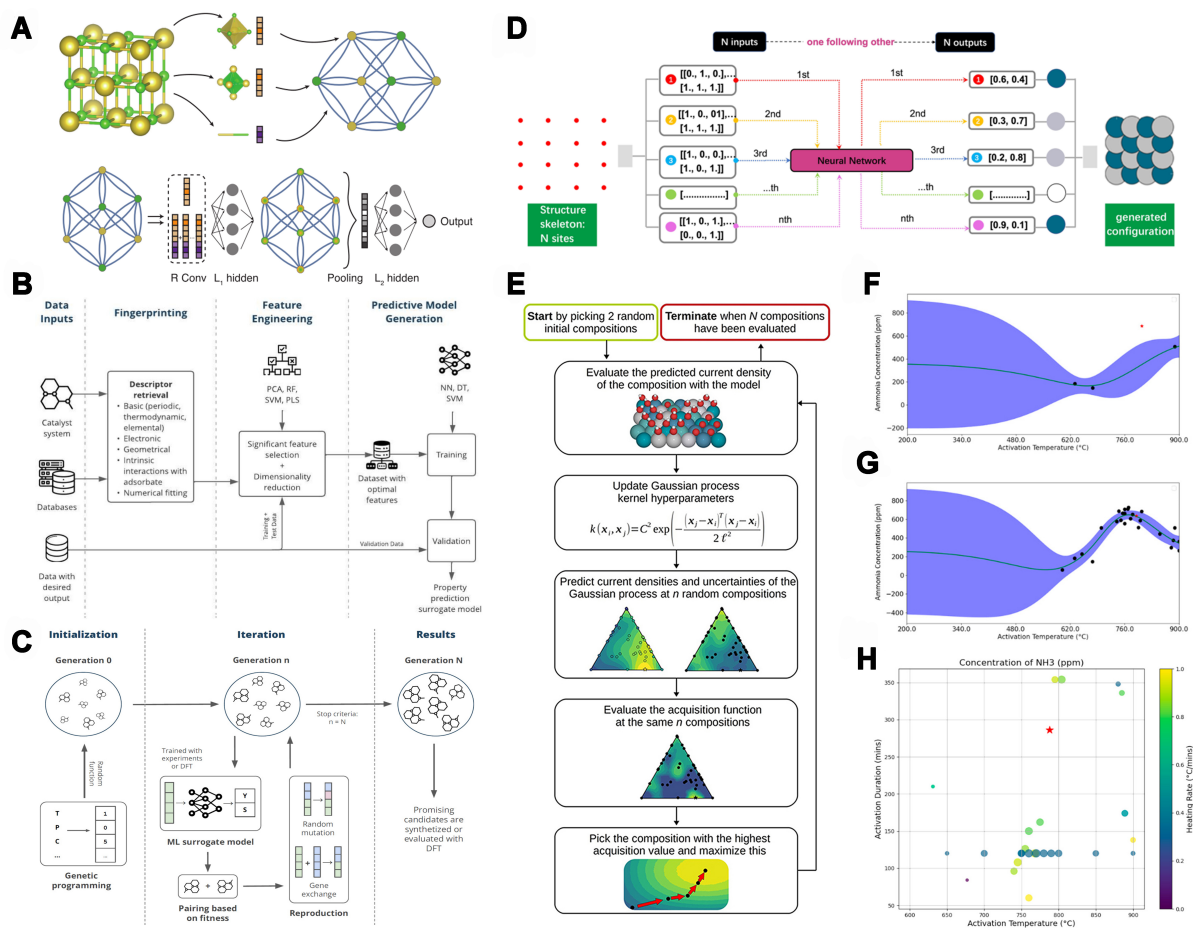


Figure 10. (A) CGCNN process diagram. CGCNN converts crystal structures into feature vectors, learning and predicting material properties. Copyright 2018, American Physical Society, Reproduced with permission^[132]; (B) Process diagram for building predictive models using ML. This performance model holds the capability to forecast catalytic-related properties based on computational data and information sourced from material databases; (C) The general steps of catalyst optimization genetic algorithm supported by AI. Copyright 2024, American Chemical Society, Reproduced with permission^[136]; (D) Roadmap for generating NES; (E) Workflow of Bayesian Optimization algorithm. Copyright 2021, OAE Publishing Inc. Reproduced with permission^[137]; (F) Initial state model with limited data points; (G) Advanced stage of optimization, model improved through a larger dataset; (H) Predicted optimal point by Bayesian optimization algorithm, along with experimental data points obtained. Copyright 2023, American Chemical Society, Reproduced with permission^[141]. CGCNN: Crystal graph convolutional neural network; ML: machine learning; AI: artificial intelligence; NES: neural evolutionary structures.

illustrate that Bayesian optimization leverages the information extracted by LLMs to approximate the unknown complex function mapping catalyst synthesis parameters to performance indicators through the construction of probabilistic models, which are continuously updated through active learning loops, ultimately converging to the global optimal solution [Figure 10H]. This workflow effectively combines knowledge extraction with practical experimentation, providing a rapid, efficient, and high-precision alternative for catalyst optimization.

All in all, the roles of generative model in SAC design include generating new structures, optimizing performance, reducing costs, and enabling inverse design. GANs show immense potential in enhancing the efficiency and precision of the entire “design-screen-optimize” process in materials science. However, their poor extrapolation capability in material design limits accurate prediction of properties outside the training data scope, especially for phenomena with unclear physical mechanisms, where these models may lack in-

depth explanations.

CONCLUSION

AI technology demonstrates comprehensive potential in the development of SACs, spanning data generation, feature analysis, and materials design. By leveraging high-throughput DFT calculations to produce extensive datasets, coupled with ML models for feature importance analysis and performance prediction, researchers can significantly accelerate the development process of SACs. With further application of NNs and generative models, AI not only facilitates the identification of high-performance catalysts but also enables the generation of new material structures tailored to specific needs. This AI-driven approach to inverse design opens up vast possibilities for innovation in SACs, promising profound impacts in fields such as electrocatalysis and energy conversion. We foresee AI applications in catalysis advancing in the following three directions:

(1) AI-Assisted Simulation of Catalytic Processes Under Realistic Environmental Conditions: AI can play a pivotal role in simulating catalytic reactions by incorporating realistic environmental factors such as temperature, pressure, and the presence of solvents or gases, along with new methods such as CPM, explicit solvation models, and Pourbaix diagram to simulate electrochemical interfaces. By integrating ML algorithms with computational methods, researchers can create more accurate models that are closer to actual reaction processes, and reflect the complexities of real-world catalytic processes, leading to better predictions of catalyst performance in practical applications and achieving targeted design and optimization of electrocatalysts.

(2) Expanding DFT Research Capabilities: Another important function of AI is to enhance DFT studies by enabling the exploration of larger systems. Currently, DFT calculations are typically limited to hundreds of atoms due to computational constraints. AI can facilitate the scaling of DFT calculations to systems containing thousands or even millions of atoms, allowing for the study of more complex materials and catalytic mechanisms. For instance, the application of MLFFs enables rapid and accurate simulations of a broader range of novel and complex multi-atomic systems that were previously challenging to model. This advancement could significantly broaden our understanding of catalysts at a fundamental level and enable the design of more efficient materials.

(3) Theoretical Models Construction and Validation: AI can assist in building comprehensive datasets that combine experimental and computational data; critical information is extracted from vast data, aiding in identifying key steps and influencing factors in electrocatalytic reactions. AI constructs universal descriptors for various systems, linking electronic structure to performance. Researchers can refine theoretical models, thereby enhancing predictive capabilities. AI delves into mechanisms of electrocatalytic reactions by simulating and predicting the processes and outcomes of these reactions. It not only serves to interpret existing experimental results but also guides the design of novel catalysts. For instance, the development of the Digital Catalysis Platform (DigCat) exemplifies the promising application of AI in electrocatalysis research.

DECLARATIONS

Authors' contributions

Methodology, writing - original draft: Yu, Q.

Conceptualization, validation, supervision, writing - review and editing: Ma, N.

Validation, writing - review and editing: Leung, C.; Liu, H.; Ren, Y.

Resources, supervision, funding acquisition: Wei, Z.

Availability of data and materials

This review article does not contain original data or materials. All relevant information is obtained from the cited literature, which is publicly available.

Financial support and sponsorship

This work was supported by the Natural Science Foundation of Xiamen City, China (3502Z202471040). The authors also acknowledge the Shenzhen Science and Technology Innovation Commission (JCYJ20220818101016034), the City University of Hong Kong (CityU 9610533), and the Shenzhen Research Institute, City University of Hong Kong. The research work described in this paper was conducted in the JC STEM Lab of Energy and Materials Physics funded by The Hong Kong Jockey Club Charities Trust.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2025.

REFERENCES

1. Cruz-Martínez, H.; Rojas-Chávez, H.; Matadamas-Ortiz, P.; et al. Current progress of Pt-based ORR electrocatalysts for PEMFCs: an integrated view combining theory and experiment. *Mater. Today. Phys.* **2021**, *19*, 100406. DOI
2. Li, C.; Tan, H.; Lin, J.; et al. Emerging Pt-based electrocatalysts with highly open nanoarchitectures for boosting oxygen reduction reaction. *Nano. Today.* **2018**, *21*, 91-105. DOI
3. Bharadwaj, N.; Das, S.; Pathak, B. Role of morphology of platinum-based nanoclusters in ORR/OER activity for nonaqueous Li-air battery applications. *ACS. Appl. Energy. Mater.* **2022**, *5*, 12561-70. DOI
4. Gasteiger, H. A.; Kocha, S. S.; Sompalli, B.; Wagner, F. T. Activity benchmarks and requirements for Pt, Pt-alloy, and non-Pt oxygen reduction catalysts for PEMFCs. *Appl. Catal. B. Environ.* **2005**, *56*, 9-35. DOI
5. Zhang, J.; Yuan, Y.; Gao, L.; Zeng, G.; Li, M.; Huang, H. Stabilizing Pt-based electrocatalysts for oxygen reduction reaction: fundamental understanding and design strategies. *Adv. Mater.* **2021**, *33*, e2006494. DOI
6. Wang, Y.; Wang, D.; Li, Y. A fundamental comprehension and recent progress in advanced Pt-based ORR nanocatalysts. *SmartMat* **2021**, *2*, 56-75. DOI
7. Sun, T.; Tang, Z.; Zang, W.; et al. Ferromagnetic single-atom spin catalyst for boosting water splitting. *Nat. Nanotechnol.* **2023**, *18*, 763-71. DOI
8. Nguyen, T. H.; Tran, D. T.; Kim, N. H.; Lee, J. H. Iron single atom-iron nanoparticles dual-deposited nitrogen-doped graphene hybrid as an innovative catalyst to enhance the oxygen reduction reaction. *Int. J. Hydrogen. Energy.* **2023**, *48*, 32294-303. DOI
9. ul, H. M.; Wu, D.; Ajmal, Z.; et al. Derived-2D Nb₄C₃T_x sheets with interfacial self-assembled Fe-N-C single-atom catalyst for electrocatalysis in water splitting and durable zinc-air battery. *Appl. Catal. B. Environ.* **2024**, *344*, 123632. DOI
10. Pan, Y.; Liu, S.; Sun, K.; et al. A bimetallic Zn/Fe polyphthalocyanine-derived single-atom Fe-N₄ catalytic site: a superior trifunctional catalyst for overall water splitting and Zn-air batteries. *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 8614-8. DOI
11. Inoue, Y. Photocatalytic water splitting by RuO₂-loaded metal oxides and nitrides with d⁰- and d¹⁰-related electronic configurations. *Energy. Environ. Sci.* **2009**, *2*, 364. DOI
12. Xiang, S.; Zhang, Z.; Wu, Z.; et al. 3D heterostructured Ti-based Bi₂MoO₆/Pd/TiO₂ photocatalysts for high-efficiency solar light driven photoelectrocatalytic hydrogen generation. *ACS. Appl. Energy. Mater.* **2019**, *2*, 558-68. DOI
13. Serrano-Ruiz, J. C.; Luque, R.; Sepúlveda-Escribano, A. Transformations of biomass-derived platform molecules: from high added-value chemicals to fuels via aqueous-phase processing. *Chem. Soc. Rev.* **2011**, *40*, 5266-81. DOI PubMed
14. Beale, A. M.; Gao, F.; Lezcano-Gonzalez, I.; Peden, C. H.; Szanyi, J. Recent advances in automotive catalysis for NO_x emission

- control by small-pore microporous materials. *Chem. Soc. Rev.* **2015**, *44*, 7371-405. [DOI](#) [PubMed](#)
15. Roy, S.; Hegde, M.; Madras, G. Catalysis for NO_x abatement. *Appl. Energy.* **2009**, *86*, 2283-97. [DOI](#)
 16. Iyyappan, J.; Gaddala, B.; Gnanasekaran, R.; Gopinath, M.; Yuvaraj, D.; Kumar, V. Critical review on wastewater treatment using photo catalytic advanced oxidation process: role of photocatalytic materials, reactor design and kinetics. *Case. Stud. Chem. Environ. Eng.* **2024**, *9*, 100599. [DOI](#)
 17. Xu, J.; Zheng, X.; Feng, Z.; et al. Organic wastewater treatment by a single-atom catalyst and electrolytically produced H₂O₂. *Nat. Sustain.* **2021**, *4*, 233-41. [DOI](#) [PubMed](#) [PMC](#)
 18. He, C.; Cheng, J.; Zhang, X.; Douthwaite, M.; Pattison, S.; Hao, Z. Recent advances in the catalytic oxidation of volatile organic compounds: a review based on pollutant sorts and sources. *Chem. Rev.* **2019**, *119*, 4471-568. [DOI](#)
 19. Kondratenko, E. V.; Mul, G.; Baltrusaitis, J.; Larrazábal, G. O.; Pérez-Ramírez, J. Status and perspectives of CO₂ conversion into fuels and chemicals by catalytic, photocatalytic and electrocatalytic processes. *Energy. Environ. Sci.* **2013**, *6*, 3112. [DOI](#)
 20. Ding, Y.; Zhang, S.; Liu, B.; Zheng, H.; Chang, C.; Ekberg, C. Recovery of precious metals from electronic waste and spent catalysts: a review. *Resour. Conserv. Recycl.* **2019**, *141*, 284-98. [DOI](#)
 21. Kim, J. H.; Shin, D.; Lee, J.; et al. A general strategy to atomically dispersed precious metal catalysts for unravelling their catalytic trends for oxygen reduction reaction. *ACS. Nano.* **2020**, *14*, 1990-2001. [DOI](#)
 22. Fang, Y.; Guo, Y. Copper-based non-precious metal heterogeneous catalysts for environmental remediation. *Chinese. J. Catal.* **2018**, *39*, 566-82. [DOI](#)
 23. Bhatt, M. D.; Lee, J. Y. Advancement of platinum (Pt)-free (non-Pt precious metals) and/or metal-free (non-precious-metals) electrocatalysts in energy applications: a review and perspectives. *Energy. Fuels.* **2020**, *34*, 6634-95. [DOI](#)
 24. Li, J.; Stephanopoulos, M. F.; Xia, Y. Introduction: heterogeneous single-atom catalysis. *Chem. Rev.* **2020**, *120*, 11699-702. [DOI](#) [PubMed](#)
 25. Yang, X. F.; Wang, A.; Qiao, B.; Li, J.; Liu, J.; Zhang, T. Single-atom catalysts: a new frontier in heterogeneous catalysis. *Acc. Chem. Res.* **2013**, *46*, 1740-8. [DOI](#)
 26. Lou, Y.; Xu, J.; Zhang, Y.; Pan, C.; Dong, Y.; Zhu, Y. Metal-support interaction for heterogeneous catalysis: from nanoparticles to single atoms. *Mater. Today. Nano.* **2020**, *12*, 100093. [DOI](#)
 27. Zhang, Y.; Yang, J.; Ge, R.; et al. The effect of coordination environment on the activity and selectivity of single-atom catalysts. *Coord. Chem. Rev.* **2022**, *461*, 214493. [DOI](#)
 28. Ji, S.; Chen, Y.; Wang, X.; Zhang, Z.; Wang, D.; Li, Y. Chemical synthesis of single atomic site catalysts. *Chem. Rev.* **2020**, *120*, 11900-55. [DOI](#)
 29. Chen, W.; Ma, B.; Zou, R. Rational design and controlled synthesis of MOF-derived single-atom catalysts. *Acc. Mater. Res.* **2025**. [DOI](#)
 30. Kan, D.; Lian, R.; Wang, D.; et al. Screening effective single-atom ORR and OER electrocatalysts from Pt decorated MXenes by first-principles calculations. *J. Mater. Chem. A.* **2020**, *8*, 17065-77. [DOI](#)
 31. Kan, D.; Wang, D.; Zhang, X.; et al. Rational design of bifunctional ORR/OER catalysts based on Pt/Pd-doped Nb₂CT₂ MXene by first-principles calculations. *J. Mater. Chem. A.* **2020**, *8*, 3097-108. [DOI](#)
 32. Zhang, X.; Xia, Z.; Li, H.; Yu, S.; Wang, S.; Sun, G. Theoretical study of the strain effect on the oxygen reduction reaction activity and stability of FeNC catalyst. *New. J. Chem.* **2020**, *44*, 6818-24. [DOI](#)
 33. Zhang, X.; Zhang, Y.; Cheng, C.; Yang, Z.; Hermansson, K. Tuning the ORR activity of Pt-based Ti₂CO₂ MXenes by varying the atomic cluster size and doping with metals. *Nanoscale* **2020**, *12*, 12497-507. [DOI](#)
 34. Wei, B.; Fu, Z.; Legut, D.; et al. Rational design of highly stable and active MXene-based bifunctional ORR/OER double-atom catalysts. *Adv. Mater.* **2021**, *33*, e2102595. [DOI](#)
 35. Zheng, J.; Sun, X.; Qiu, C.; et al. High-throughput screening of hydrogen evolution reaction catalysts in MXene materials. *J. Phys. Chem. C.* **2020**, *124*, 13695-705. [DOI](#)
 36. Ma, N.; Zhang, Y.; Wang, Y.; et al. Machine learning-assisted exploration of the intrinsic factors affecting the catalytic activity of ORR/OER bifunctional catalysts. *Appl. Surf. Sci.* **2023**, *628*, 157225. [DOI](#)
 37. Park, N. H.; Manica, M.; Born, J.; et al. Artificial intelligence driven design of catalysts and materials for ring opening polymerization using a domain-specific language. *Nat. Commun.* **2023**, *14*, 3686. [DOI](#) [PubMed](#) [PMC](#)
 38. Mazheika, A.; Geske, M.; Müller, M.; Schunk, S.; Rosowski, F.; Kraehnert, R. Data-driven design of catalytic materials in methane oxidation based on a site isolation concept. *ACS. Catal.* **2024**, *14*, 12297-309. [DOI](#)
 39. Tamtaji, M.; Gao, H.; Hossain, M. D.; et al. Machine learning for design principles for single atom catalysts towards electrochemical reactions. *J. Mater. Chem. A.* **2022**, *10*, 15309-31. [DOI](#)
 40. Sun, J.; Tu, R.; Xu, Y.; et al. Machine learning aided design of single-atom alloy catalysts for methane cracking. *Nat. Commun.* **2024**, *15*, 6036. [DOI](#) [PubMed](#) [PMC](#)
 41. Fu, H.; Li, K.; Zhang, C.; et al. Machine-learning-assisted optimization of a single-atom coordination environment for accelerated fenton catalysis. *ACS. Nano.* **2023**, *17*, 13851-60. [DOI](#)
 42. Tamtaji, M.; Chen, S.; Hu, Z.; Goddard, I. W. A.; Chen, G. A surrogate machine learning model for the design of single-atom catalyst on carbon and porphyrin supports towards electrochemistry. *J. Phys. Chem. C.* **2023**, *127*, 9992-10000. [DOI](#)
 43. Gu, G. H.; Noh, J.; Kim, S.; Back, S.; Ulissi, Z.; Jung, Y. Practical deep-learning representation for fast heterogeneous catalyst screening. *J. Phys. Chem. Lett.* **2020**, *11*, 3185-91. [DOI](#)

44. Mueller, T.; Kusne, A. G.; Ramprasad, R. Machine learning in materials science. In: Parrill AL, Lipkowitz KB, editors. Reviews in computational chemistry. Wiley; 2016. pp. 186-273. DOI
45. Raccuglia, P.; Elbert, K. C.; Adler, P. D.; et al. Machine-learning-assisted materials discovery using failed experiments. *Nature* **2016**, *533*, 73-6. DOI
46. Li, H.; Zhang, Z.; Liu, Z. Application of artificial neural networks for catalysis: a review. *Catalysts* **2017**, *7*, 306. DOI
47. Pillai, H. S.; Li, Y.; Wang, S. H.; et al. Interpretable design of Ir-free trimetallic electrocatalysts for ammonia oxidation with graph neural networks. *Nat. Commun.* **2023**, *14*, 792. DOI PubMed PMC
48. Janet, J. P.; Ramesh, S.; Duan, C.; Kulik, H. J. Accurate multiobjective design in a space of millions of transition metal complexes with neural-network-driven efficient global optimization. *ACS. Cent. Sci.* **2020**, *6*, 513-24. DOI PubMed PMC
49. Mu, Y.; Sun, L. Catalyst optimization design based on artificial neural network. *AJRCoS* **2022**, *13*, 1-12. DOI
50. Hansen, K.; Montavon, G.; Biegler, F.; et al. Assessment and validation of machine learning methods for predicting molecular atomization energies. *J. Chem. Theory. Comput.* **2013**, *9*, 3404-19. DOI PubMed
51. Mu, Y.; Wang, T.; Zhang, J.; Meng, C.; Zhang, Y.; Kou, Z. Single-atom catalysts: advances and challenges in metal-support interactions for enhanced electrocatalysis. *Electrochem. Energy. Rev.* **2022**, *5*, 145-86. DOI
52. Zhang, J.; Yang, H.; Liu, B. Coordination engineering of single-atom catalysts for the oxygen reduction reaction: a review. *Adv. Energy. Mater.* **2021**, *11*, 2002473. DOI
53. dos, S. T. C.; Mancera, R. C.; Rocha, M. V.; et al. CO₂ and H₂ adsorption on 3D nitrogen-doped porous graphene: experimental and theoretical studies. *J. CO₂ Util.* **2021**, *48*, 101517. DOI
54. Chen, B.; Li, L.; Liu, L.; Cao, J. Molecular simulation of adsorption properties of thiol-functionalized titanium dioxide (TiO₂) nanostructure for heavy metal ions removal from aqueous solution. *J. Mol. Liq.* **2022**, *346*, 118281. DOI
55. Thirumalai, H.; Kitchin, J. R. Investigating the reactivity of single atom alloys using density functional theory. *Top. Catal.* **2018**, *61*, 462-74. DOI
56. He, C.; Lee, C. H.; Meng, L.; Chen, H. T.; Li, Z. Selective orbital coupling: an adsorption mechanism in single-atom catalysis. *J. Am. Chem. Soc.* **2024**, *146*, 12395-400. DOI
57. Zhang, Y.; Wang, Y.; Ma, N.; Liang, B.; Xiong, Y.; Fan, J. Revealing the adsorption behavior of nitrogen reduction reaction on strained Ti₂CO₂ by a spin-polarized *d*-band center model. *Small* **2024**, *20*, e2306840. DOI
58. Ma, N.; Li, N.; Wang, T.; Ma, X.; Fan, J. Strain engineering in the oxygen reduction reaction and oxygen evolution reaction catalyzed by Pt-doped Ti₂CF₂. *J. Mater. Chem. A.* **2022**, *10*, 1390-401. DOI
59. Hutchison, P.; Rice, P. S.; Warburton, R. E.; Raugei, S.; Hammes-Schiffer, S. Multilevel computational studies reveal the importance of axial ligand for oxygen reduction reaction on Fe-N-C materials. *J. Am. Chem. Soc.* **2022**, *144*, 16524-34. DOI PubMed
60. Xiao, G.; Lu, R.; Liu, J.; Liao, X.; Wang, Z.; Zhao, Y. Coordination environments tune the activity of oxygen catalysis on single atom catalysts: a computational study. *Nano. Res.* **2022**, *15*, 3073-81. DOI
61. Zheng, X.; Chen, S.; Li, J.; et al. Two-dimensional carbon graphdiyne: advances in fundamental and application research. *ACS. Nano.* **2023**, *17*, 14309-46. DOI
62. Ullah, F.; Ayub, K.; Mahmood, T. High performance SACs for HER process using late first-row transition metals anchored on graphyne support: a DFT insight. *Int. J. Hydrogen. Energy.* **2021**, *46*, 37814-23. DOI
63. Liu, T.; Wang, G.; Bao, X. Electrochemical CO₂ reduction reaction on 3d transition metal single-atom catalysts supported on graphdiyne: a DFT study. *J. Phys. Chem. C.* **2021**, *125*, 26013-20. DOI
64. Li, J.; Li, K.; Li, Z.; et al. Capture of single Ag atoms through high-temperature-induced crystal plane reconstruction. *Nat. Commun.* **2024**, *15*, 3874. DOI PubMed PMC
65. Wang, Y. G.; Mei, D.; Glezakou, V. A.; Li, J.; Rousseau, R. Dynamic formation of single-atom catalytic active sites on ceria-supported gold nanoparticles. *Nat. Commun.* **2015**, *6*, 6511. DOI PubMed PMC
66. Fan, Q. Y.; Sun, J. J.; Wang, F.; Cheng, J. Adsorption-induced liquid-to-solid phase transition of Cu clusters in catalytic dissociation of CO₂. *J. Phys. Chem. Lett.* **2020**, *11*, 7954-9. DOI
67. Jia, C.; Sun, Q.; Liu, R.; et al. Challenges and opportunities for single-atom electrocatalysts: from lab-scale research to potential industry-level applications. *Adv. Mater.* **2024**, *36*, e2404659. DOI
68. Zhuo, H. Y.; Zhang, X.; Liang, J. X.; Yu, Q.; Xiao, H.; Li, J. Theoretical understandings of graphene-based metal single-atom catalysts: stability and catalytic performance. *Chem. Rev.* **2020**, *120*, 12315-41. DOI
69. Di Liberto G, Giordano L, Pacchioni G. Predicting the stability of single-atom catalysts in electrochemical reactions. *ACS. Catal.* **2024**, *14*, 45-55. DOI
70. Tan, S.; Ji, Y.; Li, Y. Single-atom electrocatalysis for hydrogen evolution based on the constant charge and constant potential models. *J. Phys. Chem. Lett.* **2022**, *13*, 7036-42. DOI
71. Cui, Y.; Ren, C.; Wu, M.; et al. Structure-stability relation of single-atom catalysts under operating conditions of CO₂ reduction. *J. Am. Chem. Soc.* **2024**, *146*, 29169-76. DOI
72. Han, Y.; Xu, H.; Li, Q.; Du, A.; Yan, X. DFT-assisted low-dimensional carbon-based electrocatalysts design and mechanism study: a review. *Front. Chem.* **2023**, *11*, 1286257. DOI PubMed PMC
73. Zhang, D.; Li, H. The potential of zero charge and solvation effects on single-atom M-N-C catalysts for oxygen electrocatalysis. *J. Mater. Chem. A.* **2024**, *12*, 13742-50. DOI
74. Liu, T.; Wang, Y.; Li, Y. How pH affects the oxygen reduction reactivity of Fe-N-C materials. *ACS. Catal.* **2023**, *13*, 1717-25. DOI

75. Unke, O. T.; Chmiela, S.; Sauceda, H. E.; et al. Machine learning force fields. *Chem. Rev.* **2021**, *121*, 10142-86. DOI PubMed PMC
76. Hu, J.; Zhou, L.; Jiang, J. Efficient machine learning force field for large-scale molecular simulations of organic systems. *CCS Chem* **2024**. DOI
77. Zhang, D.; Yi, P.; Lai, X.; Peng, L.; Li, H. Active machine learning model for the dynamic simulation and growth mechanisms of carbon on metal surface. *Nat. Commun.* **2024**, *15*, 344. DOI PubMed PMC
78. Quan, X.; Cheng, M.; Wang, K.; et al. High-throughput screening technologies of efficient catalysts for the ammonia economy. *ChemCatChem* **2025**, *17*, e202401001. DOI
79. Han, Z. K.; Sarker, D.; Ouyang, R.; Mazheika, A.; Gao, Y.; Levchenko, S. V. Single-atom alloy catalysts designed by first-principles calculations and artificial intelligence. *Nat. Commun.* **2021**, *12*, 1833. DOI PubMed PMC
80. Wang, Y.; Hu, R.; Li, Y.; Wang, F.; Shang, J.; Shui, J. High-throughput screening of carbon-supported single metal atom catalysts for oxygen reduction reaction. *Nano. Res.* **2022**, *15*, 1054-60. DOI
81. Yue, Y.; Chen, Y.; Zhang, X.; Qin, J.; Zhang, X.; Liu, R. High-throughput screening of highly active and selective single-atom catalysts for ammonia synthesis on WB2 (0 0 1) surface. *Appl. Surf. Sci.* **2022**, *606*, 154935. DOI
82. Xue, Z.; Tan, R.; Tian, J.; Hou, H.; Zhang, X.; Zhao, Y. Unraveling the activity trends of T-C₂N based single-atom catalysts for electrocatalytic nitrate reduction via high-throughput screening. *J. Colloid. Interface. Sci.* **2024**, *674*, 353-60. DOI
83. Nandy, A.; Duan, C.; Taylor, M. G.; Liu, F.; Steeves, A. H.; Kulik, H. J. Computational discovery of transition-metal complexes: from high-throughput screening to machine learning. *Chem. Rev.* **2021**, *121*, 9927-10000. DOI PubMed
84. Shen, L.; Zhou, J.; Yang, T.; Yang, M.; Feng, Y. P. High-throughput computational discovery and intelligent design of two-dimensional functional materials for various applications. *Acc. Mater. Res.* **2022**, *3*, 572-83. DOI
85. Xu, D.; Zhang, Q.; Huo, X.; Wang, Y.; Yang, M. Advances in data-assisted high-throughput computations for material design. *Mater. Genome. Eng. Adv.* **2023**, *1*, e11. DOI
86. Ma, N.; Wang, Y.; Zhang, Y.; Liang, B.; Zhao, J.; Fan, J. First-principles screening of Pt doped Ti₂CNL (N = O, S and Se, L = F, Cl, Br and I) as high-performance catalysts for ORR/OER. *Appl. Surf. Sci.* **2022**, *596*, 153574. DOI
87. Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials discovery and design using machine learning. *J. Materiom.* **2017**, *3*, 159-77. DOI
88. Wei, J.; Chu, X.; Sun, X.; et al. Machine learning in materials science. *InfoMat* **2019**, *1*, 338-58. DOI
89. Chen, H.; Zheng, Y.; Li, J.; Li, L.; Wang, X. AI for nanomaterials development in clean energy and carbon capture, utilization and storage (CCUS). *ACS. Nano.* **2023**, *17*, 9763-92. DOI
90. Chen, A.; Zhang, X.; Chen, L.; Yao, S.; Zhou, Z. A machine learning model on simple features for CO₂ reduction electrocatalysts. *J. Phys. Chem. C.* **2020**, *124*, 22471-8. DOI
91. Zhang, Y.; Wang, Y.; Ma, N.; Fan, J. Directly predicting N₂ electroreduction reaction free energy using interpretable machine learning with non-DFT calculated features. *J. Energy. Chem.* **2024**, *97*, 139-48. DOI
92. Shi, Y.; Liang, Z. Machine learning accelerates the screening of single-atom catalysts towards CO₂ electroreduction. *Appl. Catal. A. Gen.* **2024**, *676*, 119674. DOI
93. Mou, L.; Du, J.; Li, Y.; Jiang, J.; Chen, L. Effective screening descriptors of metal-organic framework-supported single-atom catalysts for electrochemical CO₂ reduction reactions: a computational study. *ACS. Catal.* **2024**, *14*, 12947-55. DOI
94. Liu, S.; Chen, Y.; Chen, C.; et al. From single-atom catalysis to dual-atom catalysis: a comprehensive review of their application in advanced oxidation processes. *Sep. Purif. Technol.* **2024**, *351*, 127989. DOI
95. Pritom, R.; Jayan, R.; Islam, M. M. Unraveling the effect of single atom catalysts on the charging behavior of nonaqueous Mg-CO₂ batteries: a combined density functional theory and machine learning approach. *J. Mater. Chem. A.* **2024**, *12*, 2335-48. DOI
96. Aklilu, E. G.; Bounahmidi, T. Machine learning applications in catalytic hydrogenation of carbon dioxide to methanol: a comprehensive review. *Int. J. Hydrogen. Energy.* **2024**, *61*, 578-602. DOI
97. Wani, A. H.; Sharma, A. Optimizing the electrocatalytic discovery with machine learning as a novel paradigm. In: Patra S, Shukla SK, Sillanpää M, editors. *Electrocatalytic materials*. Cham: Springer Nature Switzerland; 2024. pp. 247-69. DOI
98. Lan, T.; Wang, H.; An, Q. Enabling high throughput deep reinforcement learning with first principles to investigate catalytic reaction mechanisms. *Nat. Commun.* **2024**, *15*, 6281. DOI PubMed PMC
99. Ding, R.; Chen, J.; Chen, Y.; Liu, J.; Bando, Y.; Wang, X. Unlocking the potential: machine learning applications in electrocatalyst design for electrochemical hydrogen energy transformation. *Chem. Soc. Rev.* **2024**, *53*, 11390-461. DOI
100. Sun, J.; Chen, A.; Guan, J.; et al. Interpretable machine learning-assisted high-throughput screening for understanding NRR electrocatalyst performance modulation between active center and C-N coordination. *Energy. Environ. Mater.* **2024**, *7*, e12693. DOI
101. Günay M, Yıldırım R. Recent advances in knowledge discovery for heterogeneous catalysis using machine learning. *Catal. Rev.* **2021**, *63*, 120-64. DOI
102. Sinha, P.; Jyothirmai, M.; Abraham, B. M.; Singh, J. K. Machine learning driven advancements in catalysis for predicting hydrogen evolution reaction activity. *Mater. Chem. Phys.* **2024**, *326*, 129805. DOI
103. Karthikeyan, M.; Mahapatra, D. M.; Razak, A. S. A.; Abahussain, A. A.; Ethiraj, B.; Singh, L. Machine learning aided synthesis and screening of HER catalyst: present developments and prospects. *Catal. Rev.* **2024**, *66*, 997-1027. DOI
104. Choudhary, K.; Decost, B.; Chen, C.; et al. Recent advances and applications of deep learning methods in materials science. *npj. Comput. Mater.* **2022**, *8*, 734. DOI
105. Shi, M.; Mo, P.; Liu, J. Deep neural network for accurate and efficient atomistic modeling of phase change memory. *IEEE. Electron. Device. Lett.* **2020**, *41*, 365-8. DOI

106. Gao, P.; Liu, Z.; Zhang, J.; Wang, J.; Henkelman, G. A fast, low-cost and simple method for predicting atomic/inter-atomic properties by combining a low dimensional deep learning model with a fragment based graph convolutional network. *Crystals* **2022**, *12*, 1740. DOI
107. Toyao, T.; Maeno, Z.; Takakusagi, S.; Kamachi, T.; Takigawa, I.; Shimizu, K. Machine learning for catalysis informatics: recent applications and prospects. *ACS Catal.* **2020**, *10*, 2260-97. DOI
108. Wu, L.; Li, T. Machine learning enabled rational design of atomic catalysts for electrochemical reactions. *Mater. Chem. Front.* **2023**, *7*, 4445-59. DOI
109. Zafari, M.; Kumar, D.; Umer, M.; Kim, K. S. Machine learning-based high throughput screening for nitrogen fixation on boron-doped single atom catalysts. *J. Mater. Chem. A* **2020**, *8*, 5209-16. DOI
110. Xiang, S.; Huang, P.; Li, J.; et al. Solving the structure of "single-atom" catalysts using machine learning - assisted XANES analysis. *Phys. Chem. Chem. Phys.* **2022**, *24*, 5116-24. DOI
111. Gu, J.; Wang, Z.; Kuen, J.; et al. Recent advances in convolutional neural networks. *Pattern. Recognit.* **2018**, *77*, 354-77. DOI
112. Yang, H.; Zhao, J.; Wang, Q.; et al. Convolutional neural networks and volcano plots: screening and prediction of two-dimensional single-atom catalysts. *arXiv* **2024**, arXiv:2402.03876. Available online: <https://doi.org/10.48550/arXiv.2402.03876>. (accessed 2025-02-07)
113. Mitchell, S.; Parés, F.; Faust, A. D.; et al. Automated image analysis for single-atom detection in catalytic materials by transmission electron microscopy. *J. Am. Chem. Soc.* **2022**, *144*, 8018-29. DOI
114. Horwath, J. P.; Zakharov, D. N.; Mégret, R.; Stach, E. A. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. *npj. Comput. Mater.* **2020**, *6*, 363. DOI
115. Aniceto-Ocaña, P.; Marqueses-Rodríguez, J.; Perez-Omil, J. A.; Calvino, J. J.; Castillo, C. E.; Lopez-Haro, M. Direct quantitative assessment of single-atom metal sites supported on powder catalysts. *Commun. Mater.* **2024**, *5*, 652. DOI
116. Rossi, K.; Ruiz-Ferrando, A.; Akl, D. F.; et al. Quantitative description of metal center organization and interactions in single-atom catalysts. *Adv. Mater.* **2024**, *36*, e2307991. DOI
117. Devi C, Sahaaya Arul Mary S, Karthikeyan N, Varalakshmi S, Talasila V, Rama Naidu G. Graph neural network-based multiscale thermal modeling for heterogeneous materials with complex structures. *Therm. Sci. Eng. Prog.* **2024**, *55*, 102983. DOI
118. Shi, X.; Zhou, L.; Huang, Y.; Wu, Y.; Hong, Z. A review on the applications of graph neural networks in materials science at the atomic scale. *Mater. Genome. Eng. Adv.* **2024**, *2*, e50. DOI
119. Price, C. C.; Singh, A.; Frey, N. C.; Shenoy, V. B. Efficient catalyst screening using graph neural networks to predict strain effects on adsorption energy. *Sci. Adv.* **2022**, *8*, eabq5944. DOI PubMed PMC
120. Boonpalit, K.; Wongnongwa, Y.; Prommin, C.; Nutanong, S.; Namuangruk, S. Data-driven discovery of graphene-based dual-atom catalysts for hydrogen evolution reaction with graph neural network and DFT calculations. *ACS Appl. Mater. Interfaces.* **2023**, *15*, 12936-45. DOI PubMed
121. Anstine, D. M.; Isayev, O. Generative models as an emerging paradigm in the chemical sciences. *J. Am. Chem. Soc.* **2023**, *145*, 8736-50. DOI PubMed PMC
122. Türk, H.; Landini, E.; Kunkel, C.; Margraf, J. T.; Reuter, K. Assessing deep generative models in chemical composition space. *Chem. Mater.* **2022**, *34*, 9455-67. DOI
123. Suvarna, M.; Pérez-Ramírez, J. Embracing data science in catalysis research. *Nat. Catal.* **2024**, *7*, 624-35. DOI
124. Niu, Z.; Zhao, W.; Deng, H.; et al. Generative artificial intelligence for designing multi-scale hydrogen fuel cell catalyst layer nanostructures. *ACS Nano.* **2024**, *18*, 20504-17. DOI PubMed PMC
125. Park, J.; Kim, H.; Kang, Y.; Lim, Y.; Kim, J. From data to discovery: recent trends of machine learning in metal-organic frameworks. *JACS Au.* **2024**, *4*, 3727-43. DOI PubMed PMC
126. Ishikawa, A. Heterogeneous catalyst design by generative adversarial network and first-principles based microkinetics. *Sci. Rep.* **2022**, *12*, 11657. DOI PubMed PMC
127. Wang, M.; Zhu, H. Machine learning for transition-metal-based hydrogen generation electrocatalysts. *ACS Catal.* **2021**, *11*, 3930-7. DOI
128. Tempke, R.; Musho, T. Autonomous design of new chemical reactions using a variational autoencoder. *Commun. Chem.* **2022**, *5*, 40. DOI PubMed PMC
129. Vignesh, R.; Balasubramani, V.; Sridhar, T. M. Machine learning for next-generation functional materials. In: Joshi N, Kushvaha V, Madhushri P, editors. Machine learning for advanced functional materials. Singapore: Springer Nature; 2023. pp. 199-219. DOI
130. Fang, J.; Xie, M.; He, X.; et al. Machine learning accelerates the materials discovery. *Mater. Today. Commun.* **2022**, *33*, 104900. DOI
131. Kim, S.; Noh, J.; Gu, G. H.; Aspuru-Guzik, A.; Jung, Y. Generative adversarial networks for crystal structure prediction. *ACS. Cent. Sci.* **2020**, *6*, 1412-20. DOI PubMed PMC
132. Xie, T.; Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **2018**, *120*, 145301. DOI PubMed
133. Li, X.; Yang, X.; Huang, Y.; Zhang, T.; Liu, B. Supported noble-metal single atoms for heterogeneous catalysis. *Adv. Mater.* **2019**, *31*, e1902031. DOI
134. Liu, D.; He, Q.; Ding, S.; Song, L. Structural regulation and support coupling effect of single-atom catalysts for heterogeneous catalysis. *Adv. Energy. Mater.* **2020**, *10*, 2001482. DOI

135. Dmitrieva, A. P.; Fomkina, A. S.; Tracey, C. T.; et al. AI and ML for selecting viable electrocatalysts: progress and perspectives. *J. Mater. Chem. A*. **2024**, *12*, 31074-102. [DOI](#)
136. Benavides-Hernández, J.; Dumeignil, F. From characterization to discovery: artificial intelligence, machine learning and high-throughput experiments for heterogeneous catalyst design. *ACS. Catal.* **2024**, *14*, 11749-79. [DOI](#)
137. Chen, L.; Chen, Z.; Yao, X.; et al. High-entropy alloy catalysts: high-throughput and machine learning-driven design. *J. Mater. Inf.* **2022**, *2*, 19. [DOI](#)
138. Su, Y.; Wang, X.; Ye, Y.; et al. Automation and machine learning augmented by large language models in a catalysis study. *Chem. Sci.* **2024**, *15*, 12200-33. [DOI](#) [PubMed](#) [PMC](#)
139. Shields, B. J.; Stevens, J.; Li, J.; et al. Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **2021**, *590*, 89-96. [DOI](#)
140. Liu, X.; Fan, K.; Huang, X.; Ge, J.; Liu, Y.; Kang, H. Recent advances in artificial intelligence boosting materials design for electrochemical energy storage. *Chem. Eng. J.* **2024**, *490*, 151625. [DOI](#)
141. Lai, N. S.; Tew, Y. S.; Zhong, X.; et al. Artificial intelligence (AI) workflow for catalyst design and optimization. *Ind. Eng. Chem. Res.* **2023**, *62*, 17835-48. [DOI](#)