catalysts

MDPI

Review

Machine Learning-Guided Design of High Entropy Oxides for Photo-Electrocatalysis and Sustainable Energy Conversion

Muhammad Bilal [1,2,†], Muhammad Sabir [3,*,†], Niaz Wali [4], Hao Sun [5], Yakup Yildirim [6,7] and Taha Radwan [8,*]

1 School of Environmental Science and Engineering, Yangzhou University, Yangzhou 225127, China
2 School of Physical Science and Technology, Yangzhou University, Yangzhou 225009, China
3 Department of Chemical, Biological, Pharmaceutical and Environmental Sciences, Industrial Chemistry Section, University of Messina, ERIC aisbl, INSTM/CASPE, Viale F. Stagno d'Alcontres 31, 98166 Messina, Italy
4 Department of Physics, Qilu Institute of Technology, Jinan 250200, China
5 Wolfson Institute for Biomedical Research, Division of Medicine, University College London, Gower St., London WC1E 6BT, UK
6 Department of Computer Engineering, Biruni University, Istanbul 34010, Turkey
7 Mathematics Research Center, Near East University, Nicosia 99138, Cyprus
8 Department of Management Information Systems, College of Business and Economics, Qassim University, Buraydah 51452, Saudi Arabia
* Correspondence: muhammad.sabir@studenti.unime.it (M.S.); t.radwan@qu.edu.sa (T.R.)
† These authors contributed equally to this work.

Abstract

High-entropy oxides (HEOs) offer a vast compositional platform for photo-electrocatalysis and sustainable energy conversion, where multicationic disorder can tune phase stability, electronic structure, defect chemistry, light absorption, charge transport, and surface reactivity. However, the enormous number of possible compositions, local configurations, and synthesis conditions limits conventional trial-and-error discovery. This review critically evaluates how machine learning (ML) accelerates HEO design, from phase and structure classification, oxygen-vacancy prediction, machine learning interatomic potentials, and active learning to photocatalytic, electrocatalytic, thermocatalytic, and energy-storage applications. Unlike earlier surveys, it grades the evidence by separating studies that report accurate ML predictions from those in which ML-selected HEOs were synthesized and shown to perform; benchmarks representative models together with their dataset sizes and validation procedures; and draws on high-entropy alloy catalysis to show how local atomic environments shape adsorption-energy distributions, reaction barriers, and selectivity. This appraisal shows that functional validation of ML-selected HEO catalysts has so far been achieved mainly for CH_4 oxidation, the water–gas shift reaction, and the oxygen evolution reaction, whereas photocatalytic H_2 evolution and CO_2 conversion still lack prospective ML validation. Key challenges include biased datasets, incomplete synthesis metadata, inadequate representations of local disorder, and uncertain applicability domains. Finally, complementary validation categories and a minimum reporting standard are proposed to support reliable ML-guided discovery of synthesizable HEOs for sustainable energy technologies.

Keywords: high-entropy oxides; machine learning; mathematical modeling and materials informatics; active learning; catalysis; energy storage and conversion; photocatalysis; materials discovery

Check for updates

Academic Editor: Dominic Alfonso

Received: 15 September 2026
Revised: 26 October 2026
Accepted: 29 October 2026
Published: 31 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

High-entropy oxides (HEOs) have emerged as a new class of compositionally complex materials in which multiple principal cations occupy a common oxide lattice, gen-

erating extensive chemical disorder and diverse local environments [1]. The resulting variations in cation–oxygen interactions, lattice distortion, oxidation states, defects, and electronic structure provide opportunities to tune properties beyond those typically accessible in conventional oxides [2]. Since the pioneering demonstration of entropy stabilized $(Mg_{0.2}Co_{0.2}Ni_{0.2}Cu_{0.2}Zn_{0.2})$ O by Rost et al. [1], HEOs have been explored across catalysis, energy storage and conversion [3], photocatalysis, and environmental applications.

The attractiveness of HEOs is accompanied by an exceptional materials design. The number of possible elemental combinations increases rapidly with the number of constituent cations, while additional degrees of freedom arise from non-equimolar compositions, oxidation states, oxygen nonstoichiometric, crystal structures, defects, and synthesis conditions [4]. Consequently, HEO discovery involves a multidimensional composition structure property landscape rather than a simple elemental selection problem. Although configurational entropy can favor chemical mixing, phase formation and functional behavior are also governed by enthalpic interactions, ionic-size mismatch, local structural environments, defects, and processing options [5]. Exhaustive experimental exploration of this space is therefore impractical, and conventional trial-and-error strategies provide only limited coverage of the theoretically accessible compositions [6].

Machine learning (ML) offers a promising route to address this challenge by learning relationships between chemical descriptors, structures, processing conditions, and experimentally or computationally determined properties. Recent HEO studies have demonstrated ML assisted crystal structure classification, machine learning interatomic potentials for large-scale composition screening, and data-driven prediction of synthesizability [7–9]. In parallel, ML has been coupled with density functional theory (DFT) to predict oxygen vacancy formation energies and identify promising high entropy perovskites, while active learning frameworks have enabled iterative selection of new HEO compositions for experimental evaluation [10,11]. These developments indicate a transition from conventional forward prediction toward increasingly targeted exploration of unexplored HEO chemical space [11,12].

Despite this progress, HEO machine learning remains constrained by limited and heterogeneous datasets, incomplete synthesis metadata, representation of local chemical disorder, model extrapolation, and the scarcity of systematically reported negative experimental results [8,13]. Moreover, most existing studies focus on phase stability, structural descriptors, or individual physicochemical properties, whereas direct prediction of application-level performance remains comparatively underdeveloped. Therefore, the next stage of HEO informatics requires integration of curated experimental data, DFT, machine learning interatomic potentials, physically meaningful descriptors, uncertainty-aware learning, and experimental validation.

This review critically examines how machine learning can transform HEO discovery from empirical composition screening toward data-driven design. The discussion progresses from the fundamentals and multidimensional chemistry of HEOs to data curation and descriptor construction, followed by ML-guided prediction of phase stability, structure, defects, electronic properties, and composition optimization. The review further considers applications in catalysis, photocatalysis, energy storage, and environmental technologies, before evaluating the major limitations and emerging opportunities in active learning, generative AI, and autonomous experimentation. The overarching perspective is that ML should complement, rather than replace, experimental and theoretical approaches by integrating compositional information with structure, defect chemistry, physicochemical properties, and functional performance, thereby enabling more efficient prediction, screening, and experimental validation of promising HEOs.

Recent reviews [14–16] have surveyed HEO chemistry, catalytic applications, and high-entropy materials informatics. The review most closely related to the present work, by He et al. [15], gives a broad, catalysis-oriented overview of artificial-intelligence-assisted HEO design: it discusses data representation and model selection, surveys AI applications to phase stability, hydrogen production, oxygen evolution, thermal catalysis, photocatalysis, and electrocatalyst discovery, and calls for multi-objective, closed-loop integration of AI, DFT, MLIP-based sampling, operando characterization, and experimentation. Mashhadimoslem et al. [16] cover computational and ML methods for high-entropy materials in energy applications more generally. The present review is complementary to these works and adds four specific contributions. First, it grades the literature by the strength of its evidence, explicitly separating studies that demonstrate accurate ML predictions from those in which an ML-selected HEO was synthesized and shown to deliver the targeted functional performance (Section 5.5.2). Second, it benchmarks representative HEO-ML studies quantitatively, reporting model metrics together with dataset sizes, validation procedures, reaction conditions, and experimental outcomes, so that headline scores are not compared out of context. Third, it treats local chemical disorder as a distribution to be represented and sampled and connects this view with lessons from high-entropy alloy (HEA) catalysis on adsorption-energy distributions, reaction barriers, and selectivity (Sections 5.1.1 and 5.1.2). Fourth, it converts these assessments into practical recommendations: chemistry-aware data splitting, a set of complementary validation categories, and a minimum reporting standard for ML-ready HEO data (Section 6.6). The central conclusion of this appraisal is that ML-guided discovery of HEOs with experimentally demonstrated catalytic performance has so far been achieved for only a few reactions and that the near-term priority of the field is better data and validation rather than more complex algorithms.

Review Scope

The recent expansion of HEO and high-entropy-material informatics makes clear positioning essential. Several reviews now cover HEO fundamentals [14,15,17], catalytic applications [18,19], electrochemical storage [18,20], photocatalysis [21], artificial intelligence in high-entropy materials [15,16], and AI-assisted HEO catalyst design [15]. The present review is therefore intended to organize to evaluate what conditions an ML workflow can move from interpolation within previously sampled chemistry toward reliable identification of experimentally synthesizable HEOs with controlled local disorder and useful functional performance.

The evidence base was convened from foundational HEO thermodynamics and local structure studies [22,23], direct HEO-specific ML and MLIP investigations [8,9,24], materials informatics methodology, out-of-distribution (OOD) validation [25–28], active learning, self-driving laboratories, and representative functional HEO studies [11,29–31]. Special attention is given to studies that report external or chemistry-oriented validation [25,28], disclose datasets or code [32,33], incorporate negative or failed experiments [34], experimentally synthesize model-selected HEO candidates [11,24]. Random train/test splitting can substantially overestimate discovery capability when chemically similar compositions occur in both subsets [25,35]. Broader methodological examples, such as ML prediction of membrane performance from synthesis conditions [36], are considered separately from direct HEO discovery.

The literature was identified by searching the Web of Science Core Collection, Scopus, and Google Scholar, complemented by backward and forward citation tracking of key articles and by direct searches of the main publisher platforms (ACS, RSC, Elsevier, Wiley, Springer Nature, and MDPI). Search strings combined HEO terms ("high-entropy oxide", "entropy-stabilized oxide", "compositionally complex oxide", and "high-entropy

perovskite, spinel, fluorite, or rock-salt") with ML terms ("machine learning", "artificial intelligence", "interatomic potential", "graph neural network", "active learning", "Bayesian optimization", and "large language model") and with application terms (catalysis, photocatalysis, electrocatalysis, water splitting, CO_2 reduction, energy storage, and chemical looping). The search covered the period from January 2015, when entropy-stabilized oxides were first reported [1], to the final search date of 31 August 2026. Studies were included when they (i) were peer-reviewed journal articles or reviews published in English and (ii) either applied ML, MLIPs, or data-driven screening directly to HEOs and reported the model, the data source, and at least one quantitative validation result, or reported HEO functional performance together with the experimental conditions needed to interpret it as an ML target. Studies of high-entropy alloys and of general materials informatics were included only when they provide methodological or mechanistic insight that transfers directly to HEOs, such as adsorption-energy distributions, uncertainty quantification, or out-of-distribution validation. Patents and studies without a description of the model or data were not used as primary evidence of ML performance.

Figure 1 compares representative reviews and highlights the HEO-specific scope and research gap addressed here. A distinction is made between direct HEO-ML evidence and the methodological literature drawn from broader materials informatics. General ML studies are included only where they address a problem that is especially consequential for HEOs, such as graph representation of local environments [10,37], charge atomistic potentials [38,39], uncertainty calibration [40,41], literature-derived synthesis data [42,43], Bayesian optimization [44,45], or autonomous experimentation [46–48]. This separation prevents methodological possibility from being presented as demonstrated HEO capability and allows application-level claims to be evaluated according to the highest level of experimental validation achieved.

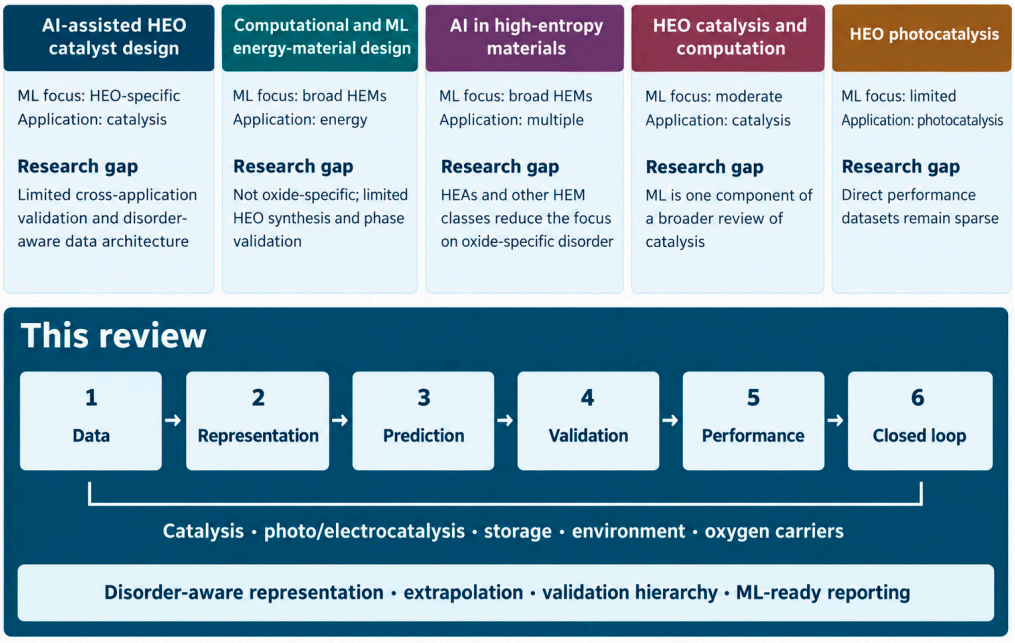


Figure 1. Scope of representative reviews and the research gap addressed by this work. The top row summarizes AI-assisted HEO catalyst design [15], computational and ML approaches to high-entropy energy materials [16], AI in high-entropy materials [13], HEO catalysis and computational insights [49], and HEO photocatalysis [21].

2. High-Entropy Oxide Fundamentals

2.1. Definition and Evolution of High-Entropy Oxides

High-entropy oxides (HEOs) represent a class of compositionally complex crystalline oxides in which multiple principal cations are incorporated into one or more crystallographic sublattices, frequently forming a solid solution phase with substantial chemical disorder [50]. The conceptual origin of HEOs can be traced to the development of high-entropy alloys (HEAs), where the incorporation of several principal metallic elements was proposed as an alternative to conventional alloy design based on one dominant element with minor additions [51,52]. The extension of this concept to oxides demonstrated that configurational complexity could similarly be exploited to access unusual combinations of structural, electronic, ionic, magnetic, and catalytic properties [49,53].

The landmark study by Rost et al. [1] introduced the concept of entropy-stabilized oxides through an equimolar rock-salt oxide, $(Mg_{0.2}Co_{0.2}Ni_{0.2}Cu_{0.2}Zn_{0.2})O$. The five cations occupy the same crystallographic sublattice, producing a compositionally complex solid solution. Importantly, the authors demonstrated that the phase behavior could be reversibly altered by temperature, providing experimental evidence that configurational entropy can contribute substantially to the stabilization of a multication oxide phase at elevated temperatures. This work established the foundation for the subsequent development of HEOs and stimulated extensive investigation of multication oxides with different compositions and crystal structures [2,4,54].

The terminology surrounding HEOs, however, requires some clarification. High entropy, chemical disorder, and entropy stabilization are related but different concepts [55]. Entropy is a thermodynamic quantity describing the number and statistical distribution of accessible configurations, whereas disorder describes the structural and chemical heterogeneity associated with those configurations. A material may therefore possess high configurational entropy without being exclusively "entropy stabilized," while a chemically disordered HEO may derive its observed phase stability from a combination of entropic and enthalpic contributions as well as kinetic effects. Kotsonis et al. [5] emphasize this difference and argue that the entropy gained through chemical mixing can favor phase formation, whereas the local disorder retained within the resulting structure can strongly influence its functional properties.

A commonly adopted practical criterion considers a high-entropy material to contain at least five principal elements distributed within a solid-solution framework, although this criterion do not represent as a universal definition [53]. In particular, the number of cations alone cannot determine whether a material should be regarded as an HEO because phase constitution, composition distribution, crystal structure, defects, and local ordering also influence its behavior. The recent HEO literature consequently regards high entropy as a design strategy for introducing extensive compositional and configurational freedom. This conceptual development has transformed HEO research from the synthesis of individual multicomponent oxides toward compositionally complex materials. The resulting materials space is substantially larger than that in conventional oxides because several cations, concentration ratios, oxidation states, structural motifs, and defect configurations can be varied simultaneously. This compositional flexibility is now recognized as one of the defining characteristics and major challenges of HEO materials design [14,53].

2.2. Configurational Entropy

The thermodynamic foundation of high-entropy materials is the configurational entropy associated with distributing different species over available crystallographic sites.

For an ideal solid solution containing n species with mole fractions x_i, the configurational entropy is given by Equation (1):

$$\Delta S_{\text{config}} = -R\sum_{i=1}^{n} x_i \ln x_i \quad (1)$$

where R is the universal gas constant and x_i represents the fraction of species i occupying the considered sublattice. For a fixed number of components, the configurational entropy reaches its maximum when the species are distributed equimolarly because this condition maximizes the number of possible distinguishable configurations.

For an equimolar five-cation system, the mole fractions are given by Equation (2):

$$x_1 = x_2 = x_3 = x_4 = x_5 = 0.2 \quad (2)$$

Substitution into Equation (1) gives Equation (3):

$$\Delta S_{\text{config}} = -R[5(0.2)\ \ln\ (0.2)] = R\ \ln\ (5) \approx 1.609R \quad (3)$$

Thus, $1.609R$ represents the ideal configurational entropy of five equimolar species occupying an equivalent sublattice. This value has consequently become a widely used benchmark for identifying high-entropy compositions. Nevertheless, $1.609R$ should not be interpreted as the experimentally measured total entropy of a real five-cation HEO. The ideal expression assumes random occupation of equivalent sites and neglects nonideal interactions and additional entropy contributions. Real HEOs can contain multiple crystallographic sublattices, oxygen vacancies, short-range chemical correlations, local structural distortions, and variations in cation oxidation states [55]. Therefore, the actual thermodynamic entropy of an HEO can deviate from the ideal configurational estimate. The difference becomes particularly important for multicomponent oxides containing both cation and anion disorder or substantial defect populations. Equimolar compositions maximize the ideal configurational entropy for a given number of components, but equimolarity is not a prerequisite for HEO formation. Non-equimolar compositions can retain substantial configurational entropy while providing additional opportunities to tune crystal structure, oxidation-state balance, lattice distortion, defect concentration, and functional properties [56]. Accordingly, restricting HEO design to equimolar compositions would unnecessarily constrain the accessible materials space [57].

The frequently used classification into low, medium, and high-entropy terms provides a convenient descriptive framework. More importantly, a high calculated configurational entropy does not guarantee the formation of a single-phase HEO [58]. Entropy contributes to the free-energy balance, but phase formation additionally depends on enthalpic interactions, ionic-size compatibility, oxidation-state requirements, structural preferences, defects, and synthesis kinetics [14]. Therefore, it can be expressed as high configurational entropy increases the thermodynamic tendency toward chemical mixing, but it does not independently determine HEO phase formation or stability. This difference is fundamental for understanding HEOs and becomes particularly important when developing predictive models, because a machine learning model trained only on configurational entropy is unlikely to capture the full complexity of HEO phase formation [58].

2.3. Compositional and Structural Diversity of HEOs

An important feature of HEO chemistry is that configurational complexity varies. Multicationic disorder has been incorporated into several oxide structural families, including rock-salt, fluorite, spinel, perovskite, and garnet structures, as illustrated in Figure 2. Each structural family provides different crystallographic sites, coordination environments, defect mechanisms, and chemical constraints [50,59].

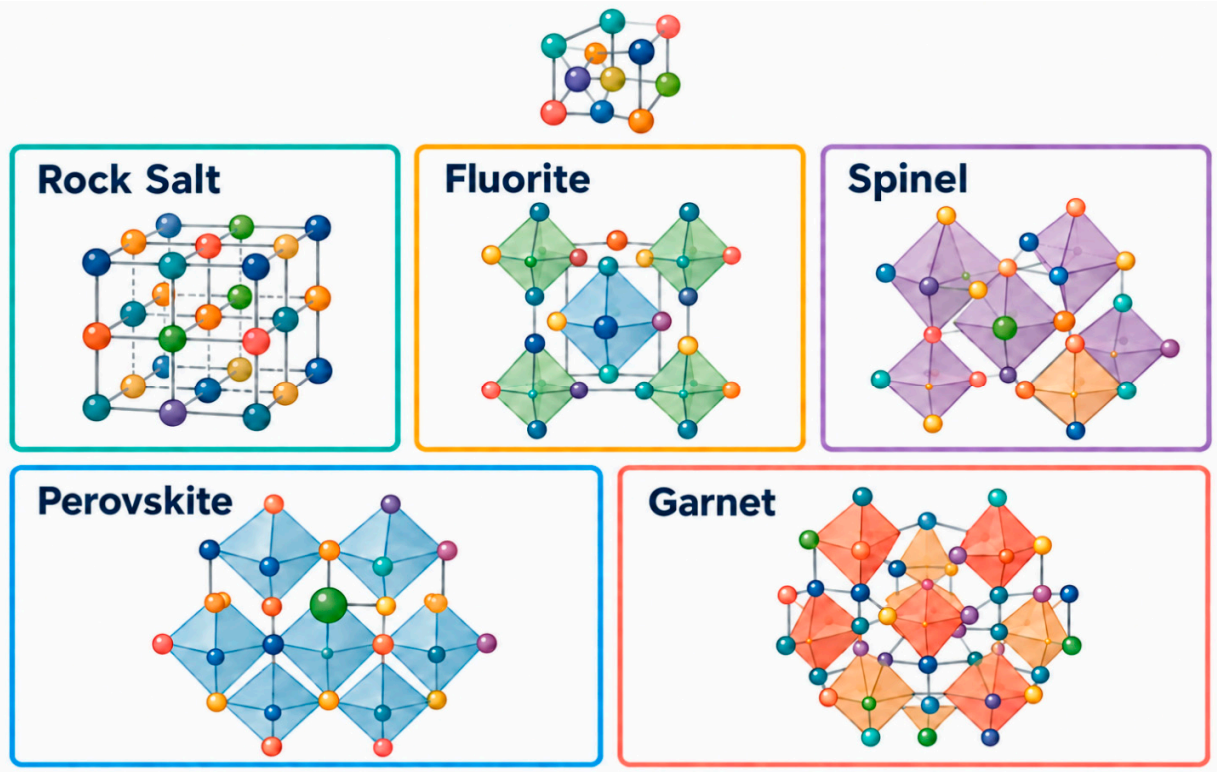


Figure 2. Representative crystal structures of high-entropy oxides: rock-salt, fluorite, spinel, perovskite, and garnet.

The rock-salt structure represents the historical starting point for HEO research. In the prototype (MgCoNiCuZn)O, multiple cations occupy a common cation sublattice while oxygen occupies the anion sublattice. Its relatively simple structure makes rock-salt HEOs valuable model systems for studying configurational entropy, cation disorder, lattice distortion, defect chemistry, and thermodynamic phase stability [1]. Fluorite HEOs provide another important structural family, particularly because their crystal chemistry can accommodate substantial oxygen non-stoichiometry. Multication substitution can modify oxygen-vacancy populations, local coordination environments, and ionic transport. The fluorite framework therefore provides an important platform for investigating the coupling between compositional disorder and defect-mediated properties [60]. In spinel HEOs, compositional complexity is accompanied by the additional possibility of distributing cations between tetrahedral and octahedral sites. Consequently, cation-site preference and local coordination become important variables in addition to overall composition. This additional degree of structural freedom can influence magnetic, catalytic, and electrochemical behavior [61]. Perovskite HEOs offer an even broader compositional design space because multiple elements can occupy different crystallographic sublattices. Changes in cation composition can modify oxidation states, coordination environments, octahedral distortions, oxygen stoichiometry, and local bonding. These characteristics make high-entropy perovskites attractive for a range of functional applications [62–64].

The existence of these diverse structural families demonstrates that HEO design cannot be reduced to simply increasing the number of constituent elements. The crystal structure itself defines the available sites and coordination environments into which the constituent elements must be accommodated. Consequently, the same combination of elements may exhibit different phase behavior depending on the targeted structural framework. This creates an important hierarchical design problem in which composition, structure, site occupancy, defects, and local chemical environments are strongly coupled. Experimental

studies on chemically versatile high-entropy pyrochlores further show that the same high-entropy design concept can be accommodated through very different coordination networks and defect chemistries, reinforcing the need for structure-aware descriptors [65].

2.4. Thermodynamic Principles Governing HEO Formation

The stability of a multication oxide can be considered from the Gibbs free energy relationship in Equation (4):

$$\Delta G_{\mathrm{mix}} = \Delta H_{\mathrm{mix}} - T\Delta S_{\mathrm{mix}} \quad (4)$$

where ΔH_{mix} represents the enthalpic contribution to mixing, T is temperature, and ΔS_{mix} represents the total entropy contribution. In an idealized description, configurational entropy is often considered the dominant entropy term, giving the approximation in Equation (5):

$$\Delta G_{\mathrm{mix}} \approx \Delta H_{\mathrm{mix}} - T\Delta S_{\mathrm{config}} \quad (5)$$

The configurational entropy term therefore becomes increasingly important with increasing temperature as illustrated in Figure 3. This provides the thermodynamic basis for the experimentally observed entropy stabilization of the prototype rock-salt HEO reported by Rost et al. [1]. At elevated temperature, the entropy contribution can lower the free energy of the chemically mixed state sufficiently to favor a single-phase solid solution over competing phase-separated states. However, the free-energy landscape of a real HEO is considerably more complicated than the configurational-entropy approximation in Equation (5). Chemically distinct cations can interact differently with one another and with oxygen, producing nonzero and composition-dependent mixing enthalpies. Strongly unfavorable interactions may promote segregation or ordering, whereas favorable interactions may facilitate incorporation into a common lattice. Structural strain arising from ionic-size differences and changes in coordination environment can also modify the enthalpic contribution [66].

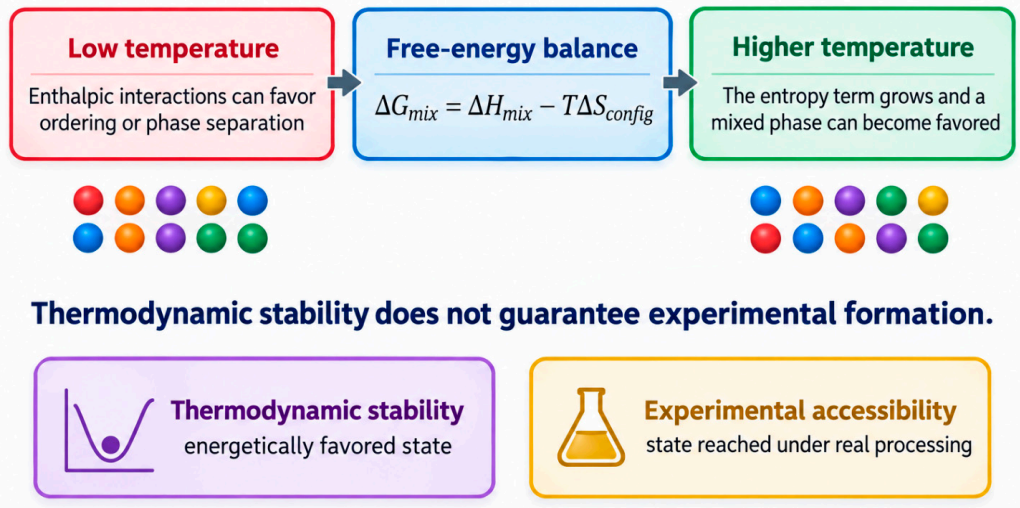


Figure 3. Schematic showing how configurational entropy can favor formation of a chemically mixed solid solution at higher temperature, while experimental phase formation also depends on synthesis accessibility.

An additional complication is that experimentally synthesized HEOs do not necessarily reach complete thermodynamic equilibrium. Rapid heating, short reaction times, quenching, nanoscale dimensions, and precursor-dependent reaction pathways can kinetically stabilize configurations that differ from the equilibrium state. Kotsonis et al. [5] emphasize that synthesis conditions can influence whether the final material contains a solid solution,

local chemical heterogeneity, or multiphase assemblies. Thus, thermodynamic stability and experimental synthesizability should be distinguished. A composition predicted to be thermodynamically favorable may not form under a particular synthesis condition, while a metastable phase may be experimentally accessible because kinetic barriers prevent transformation into the equilibrium assemblage. This distinction is particularly important for data-driven HEO discovery because a model predicting thermodynamic stability does not necessarily predict whether the corresponding composition can be synthesized experimentally under realistic conditions.

2.5. Lattice Distortion and Local Chemical Environments

The incorporation of multiple chemically distinct cations into a common lattice inevitably produces a distribution of local chemical environments. Although conventional XRD may reveal a single average crystal structure, the local environment surrounding individual cations can vary substantially in bond length, coordination geometry, oxidation state, and bonding character [15]. Recent studies have therefore shifted the description of HEOs from an idealized "random solid solution" toward a more nuanced picture involving multiscale structural and chemical heterogeneity. Cation-size mismatch is one important source of local lattice distortion. When cations with different ionic radii occupy nominally equivalent crystallographic sites, the surrounding oxygen framework must locally adapt to different bonding requirements. This can produce variations in bond lengths, local strain, coordination geometry, and lattice displacement. Similarly, differences in cation valence and electronic configuration can modify local bonding and coordination preferences. Figure 4 illustrates the high-entropy effect, lattice distortion, sluggish diffusion, and cocktail effect commonly discussed in high-entropy materials.

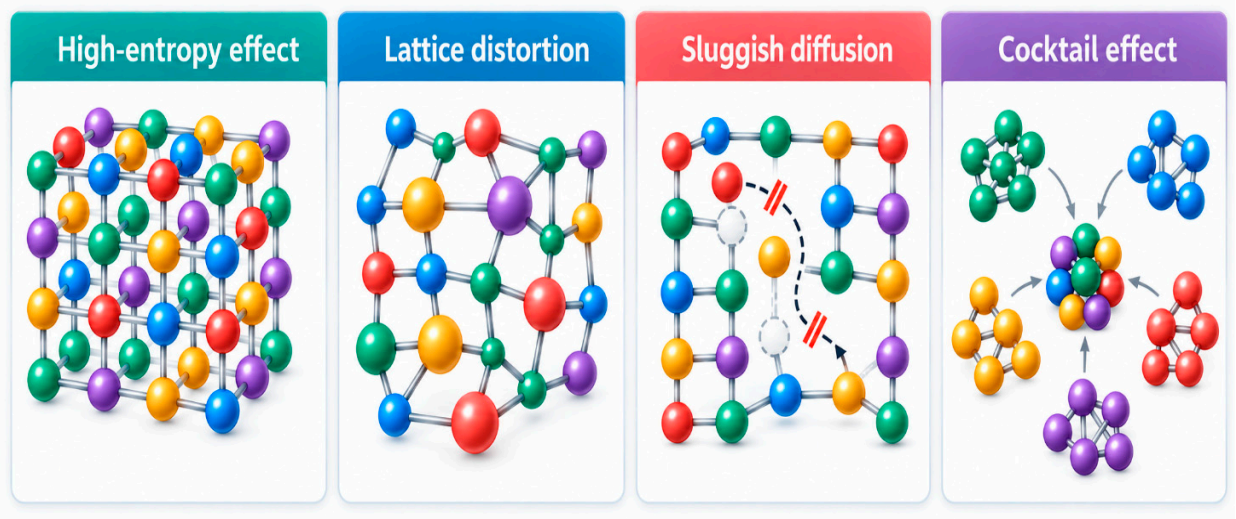


Figure 4. Schematic illustration of four effects commonly associated with high-entropy materials: high-entropy effect, lattice distortion, sluggish diffusion, and cocktail effect.

Importantly, lattice distortion does not act as a single uniform strain parameter. Instead, HEOs can contain a distribution of local distortions, reflecting the large number of possible combinations surrounding any given lattice site. It is important to emphasis that local cation-size and valence mismatch, oxygen-sublattice distortions, and site-specific bonding can strongly influence transport, magnetic, dielectric, and other functional properties [5]. This local heterogeneity provides an important distinction between conventional and high-entropy oxides. In conventional oxide, chemically equivalent sites often have relatively similar local environments, whereas an HEO can contain many chemically distinct but

crystallographically equivalent sites. Consequently, an HEO may be crystallographically ordered at the long-range scale while simultaneously exhibiting substantial local structural and chemical disorder [67].

This difference has major implications for HEO design. The macroscopic composition alone does not uniquely define the microscopic state of an HEO because different local arrangements can exist at the same overall composition. Therefore, two samples having nominally identical chemical formulas may exhibit different local ordering, defect distributions, or functional behavior depending on synthesis history. For this reason, advanced characterization methods including total scattering, spectroscopy, and high-resolution microscopy have become increasingly important for resolving short- and intermediate-range order that is not readily captured by conventional diffraction techniques.

2.6. Defects and Oxygen-Vacancy Engineering

Defect chemistry represents another fundamental dimension of HEO design. The simultaneous presence of multiple cations with different valence states and coordination preferences can generate complex charge-compensation mechanisms. In particular, oxygen vacancies can become coupled to cation oxidation-state changes, local structural relaxation, and changes in electronic structure [68]. Oxygen vacancies are especially important in oxide systems because they influence oxygen mobility, redox behavior, electronic conductivity, adsorption characteristics, and catalytic activity. Their concentration and spatial distribution can be modified through cation selection, synthesis atmosphere, temperature, and oxygen chemical potential. Nevertheless, oxygen vacancies should not be regarded universally as beneficial defects. Their functional effect depends on their concentration, charge state, local coordination, and interaction with neighboring cations. A moderate concentration of oxygen vacancies may provide favorable active sites or facilitate charge transport, whereas excessive defect concentrations can destabilize the lattice or introduce undesirable electronic states [68,69].

Similarly, cation vacancies and mixed-valence states can contribute to charge compensation and local structural relaxation. The resulting defect landscape is therefore significantly more complex than in many conventional single-cation oxides. The importance of defect chemistry further increases the dimensionality of the HEO design space. Two HEOs with identical nominal cation compositions can possess different oxygen-vacancy concentrations and oxidation-state distributions because of differences in synthesis temperature, atmosphere, precursor chemistry, or cooling rate. Hence, composition alone cannot always predict defect structure or functional performance. This observation is highly relevant to machine learning because defect-related descriptors may ultimately be required to distinguish chemically similar HEOs with substantially different properties. Figure 4 summarizes four effects commonly discussed in high-entropy materials; their relevance and magnitude depend on the particular oxide system.

2.7. Electronic and Physicochemical Characteristics

The chemical complexity of HEOs provides multiple pathways for modifying their electronic and physicochemical characteristics. Different cations introduce distinct electronic configurations, oxidation states, orbital energies, electronegativities, and bonding strengths, while local lattice distortion and defects further modify the electronic environment [9]. The resulting material can therefore exhibit an electronic structure that differs substantially from a simple linear combination of the properties of its constituent oxides. Cation substitution can modify band structure, carrier concentration, charge distribution, and redox chemistry, whereas oxygen vacancies can introduce defect-associated electronic states and alter the local oxidation-state distribution. These effects provide opportunities

for tuning optical, electrical, magnetic, electrochemical, and catalytic behavior [18,70]. An important consequence of multication disorder is the creation of chemically heterogeneous local environments. Different surface sites may contain different combinations of neighboring cations and oxygen configurations, resulting in a distribution of adsorption energies and reaction pathways. This characteristic is particularly attractive for catalytic applications because it can provide multiple chemically distinct active environments within a nominally single-phase material. However, the phrase "cocktail effect" should be used cautiously. The enhanced properties of an HEO cannot automatically be attributed to an unspecified synergistic effect of all constituent elements. Instead, the observed behavior should ideally be connected to identifiable structural, electronic, defect, or interfacial mechanisms [70,71].

Accordingly, HEOs are a high-dimensional platform, in which composition, local environment, defects, and structural framework can be independently or cooperatively manipulated. This makes HEOs attractive for applications including catalysis, energy storage, electrochemical conversion, dielectric materials, magnetic materials, and other functional technologies.

2.8. Synthesis and Characterization as Design Variables

The synthesis of HEOs requires efficient mixing of multiple cations and appropriate thermal or chemical conditions for stabilizing the desired phase [72]. Conventional solid-state synthesis is one of the most widely used approaches because of its simplicity and scalability. In this method, oxide, carbonate, nitrate, or other precursor powders are mixed and subjected to high-temperature treatment. However, the relatively long diffusion distances can lead to incomplete homogenization, and high temperatures may promote particle coarsening or unwanted phase formation [73].

Sol–gel and coprecipitation methods can provide improved elemental homogeneity because the constituent cations are mixed at the molecular or nanoscale before calcination. These approaches can reduce diffusion limitations and facilitate formation of homogeneous HEO phases at lower temperatures than conventional solid-state reactions [20]. Hydrothermal and solvothermal methods offer additional control over nucleation, crystal growth, morphology, and crystallinity. These approaches can be particularly advantageous for nanoscale HEOs and phases that are difficult to obtain through conventional high-temperature reactions. Combustion, spray pyrolysis, microwave-assisted, and other solution-based approaches can further reduce reaction times and provide control over particle size and morphology. Template-assisted approaches can additionally introduce morphological control, while thin-film deposition methods can be used to prepare HEO coatings and model systems [20,74]. It can affect phase constitution, crystallite size, local chemical ordering, oxygen-vacancy concentration, oxidation state, surface area, and defect population. Consequently, synthesis conditions constitute an important part of the HEO design space rather than an independent variable that can be ignored when relating composition to properties.

Related plasma-assisted nanoparticle studies illustrate why synthesis descriptors should accompany composition data. Ismail et al. used an atmospheric-pressure plasma–liquid process to synthesize fructose-stabilized Pd nanoparticles and related discharge-current-dependent plasma characteristics to particle formation and catalytic pollutant reduction [75]. Wali et al. reviewed how precursor-reduction kinetics, reaction–transport balance, atomic mixing, and surface segregation affect plasma-assisted trimetallic nanoparticles, emphasizing that nominal composition alone does not establish a homogeneous nanoalloy [76]. These non-HEO studies provide methodological context for recording discharge conditions, precursor chemistry, treatment time, and local characterization; they do not demonstrate plasma synthesis of an ML-selected HEO.

Characterization of HEOs similarly requires a combination of complementary techniques. X-ray diffraction (XRD) is fundamental for identifying the average long-range crystal structure and detecting crystalline secondary phases. However, because XRD represents an average structural response, it may not reveal substantial local chemical heterogeneity or short-range ordering. Transmission electron microscopy (TEM), high-resolution TEM, scanning TEM, and energy-dispersive X-ray spectroscopy (EDS) provide information on morphology, crystallinity, and nanoscale elemental distributions. X-ray photoelectron spectroscopy (XPS) provides surface-sensitive information on elemental composition and oxidation states, whereas X-ray absorption spectroscopy (XAS) can probe local coordination and oxidation-state environments. Electron paramagnetic resonance (EPR) is particularly valuable for investigating paramagnetic defect centers, while Raman and FTIR spectroscopy provide information on local bonding and vibrational environments. More advanced techniques such as total-scattering/PDF analysis and synchrotron-based spectroscopy can reveal short- and intermediate-range structural correlations that remain inaccessible to conventional XRD. The increasing use of these complementary techniques reflects an important conceptual development in HEO research: average crystallographic structure is insufficient to fully describe a chemically complex oxide. Understanding HEO properties increasingly requires information spanning multiple structural length scales.

2.9. HEO Chemistry and Machine Learning

Figure 5 summarizes a data-driven HEO discovery workflow integrating data sources, descriptor construction, ML modeling, and experimental validation. For conventional composition-based ML, material is often represented by a single vector of elemental fractions and composition-weighted statistics. For HEOs, this representation is frequently non-unique because one nominal composition can correspond to many cation arrangements, oxidation-state distributions, vacancy configurations, short-range-order motifs, and sublattice occupancies [57,59,61,66]. Local structural studies using EXAFS, total scattering/PDF [22,77], atomic-resolution STEM, EELS [78,79], Mössbauer spectroscopy, X-ray magnetic circular dichroism, and atom-probe tomography [23,80,81] have repeatedly shown that crystallographic single-phase character does not imply chemically equivalent local environments. Recent spinel studies are particularly instructive because experimentally resolved tetrahedral and octahedral occupancies depart substantially from the equal-site distributions often assumed when ideal configurational entropy is calculated [61,66].

This has direct consequences for descriptor design. Composition-averaged ionic radii, electronegativity, and configurational entropy remain useful low-cost descriptors, but they cannot distinguish samples having the same nominal formula and different site occupancies, oxygen non-stoichiometry, short-range ordering, or synthesis-dependent valence states [53,82]. A more realistic representation should therefore be hierarchical: global composition and processing variables define an ensemble of accessible local configurations [35,53,83]; local structure and defects determine electronic and surface states [57,84], and these states govern the experimentally measured functional response [85,86]. This hierarchy also explains why composition-only models can perform well for interpolation while failing when transferred to new structural families or defect regimes.

For disordered oxides, promising routes include special quasirandom structures, configurational ensembles, cluster expansion, Monte Carlo sampling [57,83], local-environment histograms [87–89], equivariant graph neural networks [90–92], and MLIPs capable of sampling many chemically distinct configurations at tractable computational cost [9,93]. The central challenge is not only model accuracy within the sampled configuration space but whether the representation and training data span the local environments encountered during synthesis, redox cycling, oxygen loss, surface reconstruction, and operation [22,28,84,86].

Consequently, HEO informatics should increasingly treat local disorder as a distribution to be represented and sampled rather than as residual noise around an average crystal structure.

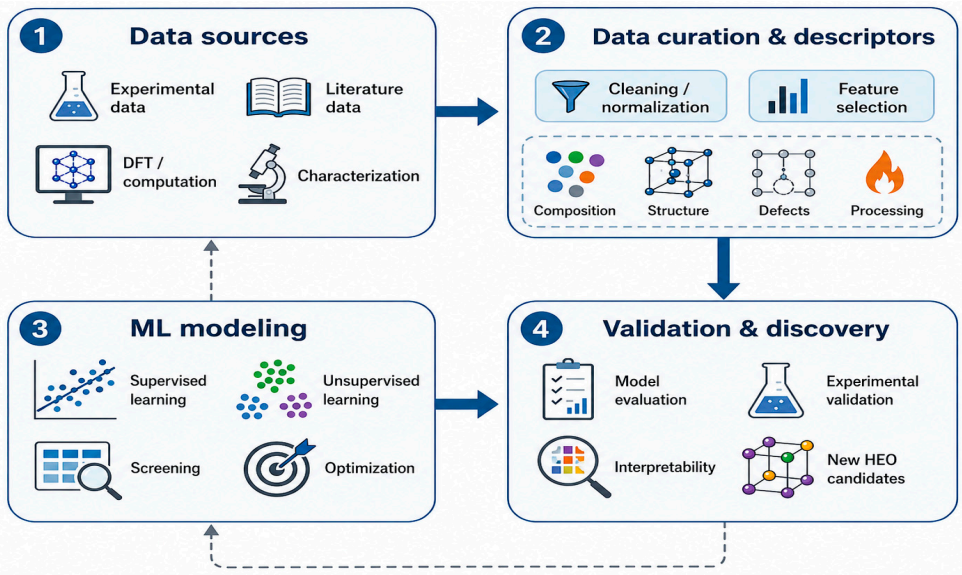


Figure 5. Schematic workflow for data-driven HEO discovery, integrating data sources, data curation and descriptor construction, machine learning modeling, and experimental validation for the identification of new HEO candidates.

3. HEO Design Through Machine Learning

The enormous compositional and structural space of high-entropy oxides (HEOs) creates a fundamental challenge for conventional materials discovery. Figure 6 illustrates how nominal composition and local structural states influence electronic, transport, and functional properties. As discussed earlier, the behavior of an HEO is governed by multiple factors such as concentration of its constituent cations but also by their oxidation states, local coordination environments, lattice distortion, defect chemistry, crystal structure, and synthesis procedure. These variables are strongly coupled, making it difficult to establish universal composition–structure–property relationships using conventional empirical approaches. Machine learning (ML) offers a complementary strategy by converting chemical, structural, thermodynamic, computational, and experimental information into quantitative descriptors that can be analyzed simultaneously. The objective is to complement experimental investigation by identifying hidden relationships within existing data, prioritize unexplored compositions, and establish a more efficient feedback loop between computation and experiment. Recent HEO studies have demonstrated this transition particularly clearly through ML-assisted phase prediction, ML interatomic potential, large-scale composition screening, and experimentally validated candidate discovery [7–9,24,94].

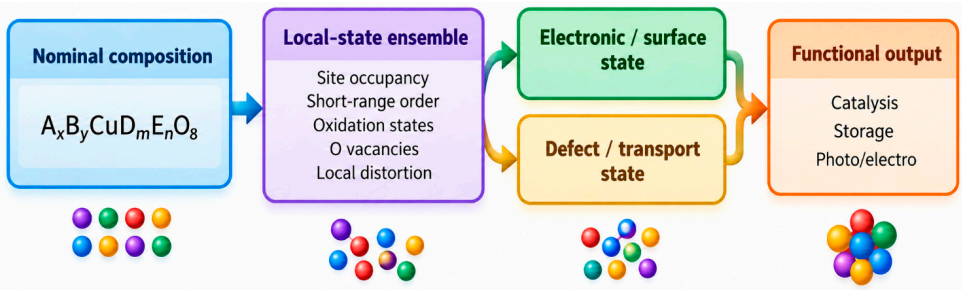


Figure 6. Schematic linking multicomponent composition and local structural states to electronic, transport, and functional properties.

3.1. From Experimental Databases to HEO Datasets

The effectiveness of an ML model is fundamentally determined by the quality, diversity, and physical relevance of the data used for training. For HEOs, useful datasets can originate from several complementary sources, including experimentally reported compositions and phase outcomes, density functional theory (DFT) calculations, high-throughput computational screening, established materials databases [95], and increasingly automated literature-mining approaches. Each source provides different types of information and introduces different uncertainties. Therefore, combining them requires careful consideration of their compatibility [96].

The experimental literature represents the most direct source of information because it contains the actual relationship between composition, processing conditions, phase constitution, and experimentally measured properties. However, HEO information is often distributed across individual publications in inconsistent formats. Chemical formulas may be reported using different conventions, synthesis conditions may be incompletely specified, and phase labels may depend on characterization techniques and detection limits. Consequently, simply aggregating literature data without systematic curation can introduce substantial noise and bias. Therefore, rigorous data curation is essential for reliable materials informatics; standardized curation practices are also necessary for reproducibility and for the development of reliable AI/ML models. Early HEO phase-mapping efforts based on empirical biplots further illustrate both the value and limitations of compact descriptor spaces, and useful separation of phase fields can be obtained from composition statistics, but such maps remain highly dependent on the coverage and labeling quality of the underlying literature data [97].

A particularly important recent development is the use of automated literature and large-language-model-based data extraction. Santos et al. [98] demonstrated an end-to-end HEO informatics workflow in which large language models were used to extract compositions, phase information, and synthesis methods from unstructured scientific abstracts, subsequently generating a structured database for ML analysis. Their reported extraction accuracy reached approximately 96% for the database-generation task, demonstrating the potential of automated literature mining to address the lack of structured HEO datasets [98–100].

Although experimental datasets provide the greatest practical relevance, their size is often insufficient for training complex ML models. DFT and high-throughput calculations can partially address this limitation by generating systematically calculated properties for large numbers of hypothetical compositions and structures. Such computational datasets can provide formation energies, electronic properties, structural parameters, defect energetics, and other descriptors that are difficult to obtain experimentally for every composition. However, DFT-derived data are not so reliable because calculated properties depend on the selected exchange correlation functional, magnetic configuration, structural model, defect configuration, and computational settings.

ML interatomic potentials (MLIPs) provide another important route for expanding HEO datasets. Unlike conventional ML models that predict a target property from predefined descriptors, MLIPs learn potential-energy surfaces from atomistic calculations and can therefore be used to relax and evaluate large numbers of candidate structures at substantially lower computational cost than direct DFT calculations. Sivak et al. [9] demonstrated this approach for rock-salt HEOs, using an ML interatomic potential to rapidly map composition space and evaluate single-phase stability using bond-length distributions and mixing enthalpy. Their approach correctly resolved experimentally stabilized compositions and identified additional candidate compositions for future investigation [8,9].

Dicks et al. [8] subsequently expanded this concept to a substantially larger HEO search space. Their workflow generated random multication structures, relaxed them using an ML interatomic potential, and evaluated candidates using descriptors based on cation-energy variance and mixing enthalpy. The methodology successfully recovered the known stable four-component α-PbO_2-type HEO and predicted additional five-component candidate systems while expanding the chemical search from seven to fourteen elements and considering three structural families [8]; related MLIP developments for disordered multicomponent oxides are discussed in [93,101,102].

3.2. Data Curation and Preprocessing

Before model construction, HEO datasets must undergo systematic curation and preprocessing. Figure 7 summarizes the normalization, curation, uncertainty assessment, and data-splitting steps required to prepare an ML-ready HEO dataset. This step is particularly important because HEO datasets are typically assembled from heterogeneous experimental and computational sources rather than generated under a single standardized protocol. Variations in composition notation, phase classification, measurement conditions, synthesis routes, and reporting practices can introduce inconsistencies that an ML model may interpret incorrectly as meaningful chemical relationships.

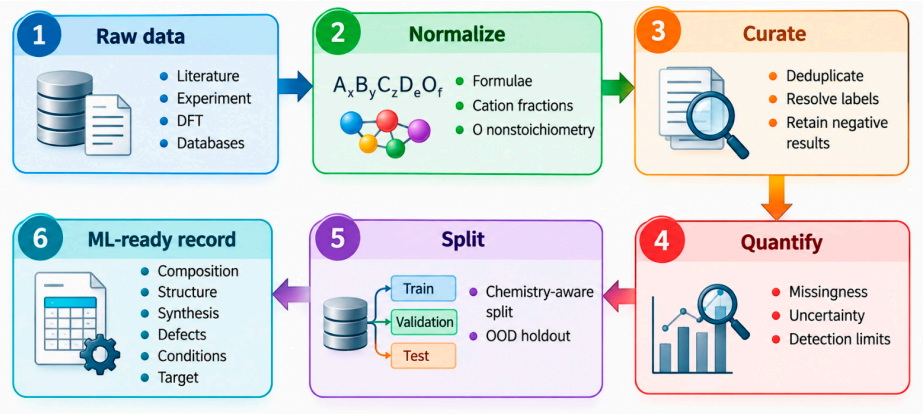


Figure 7. Workflow for converting raw materials data into an ML-ready dataset through normalization, curation, uncertainty assessment, chemistry-aware data splitting, and structured feature preparation.

The first requirement is composition normalization. Equivalent compositions may be reported using different formula conventions, while non-equimolar HEOs require careful conversion of nominal stoichiometric coefficients into consistent cation fractions. Oxygen non-stoichiometry should also be retained rather than automatically forcing all compositions into an idealized ABOx representation. Similarly, oxidation-state information is separated from elemental identity whenever it is available because the same element can contribute to very different chemical behavior depending on its oxidation state [35,103].

Duplicate identification is another important consideration. Multiple publications may report the same composition, or the same experimental sample may appear in several databases. Treating duplicate records as independent observations artificially increases the effective dataset size and can lead to overly optimistic model performance. Related compositions should also be examined carefully because random splitting can place highly similar compositions in both training and test datasets, thereby making the prediction task appear easier than genuine extrapolation into unexplored chemical space.

Missing information presents a further challenge. For example, synthesis temperature may be reported in one study but omitted in another. Similarly, oxygen-vacancy concentration or oxidation-state distributions may not be available for all samples. Missing values

should therefore be taken with caution. Depending on the model and dataset, appropriate approaches include physically informed imputation, missing data indicators, exclusion of variables with excessive missing data, or models capable of handling incomplete information. Outlier detection is equally important but should be performed cautiously. An apparent outlier may represent an experimental error, an incorrectly digitized value, or a genuinely unusual HEO composition. Removing unusual data solely because they reduce model performance risks eliminating precisely the compositions that provide information about the limits of a design space. Uncertainty should therefore be taken as an intrinsic component of HEO data. Experimental uncertainty may arise from phase identification, compositional deviations, instrument sensitivity, particle-size effects, or variations in synthesis conditions, whereas computational uncertainty can arise from the choice of DFT methodology or structural model. Hart et al. [33] emphasized that rigorous data-curation procedures improve the reliability, reproducibility, and integrity of materials datasets and are essential for reliable AI/ML applications.

3.3. Chemical Descriptors for HEOs

The main challenge in HEO machine learning is the conversion of chemical complexity into numerical representations that retain physically meaningful information. A descriptor is mathematical input variable to encode a particular aspect of HEO chemistry into a form that an ML algorithm can analyze. For HEOs, descriptors can be organized into six interconnected categories like elemental, thermodynamic, structural, electronic, defect-related, and synthesis descriptors.

3.3.1. Elemental Descriptors

Elemental descriptors represent the intrinsic chemical characteristics of the constituent cations. Common descriptors include atomic number, atomic mass, atomic radius, ionic radius, electronegativity, electron affinity, ionization energy, valence, and electronic configuration. However, using the value for each element independently may be insufficient for HEOs because several cations simultaneously occupy the same crystallographic sublattice. Consequently, composition weighted statistical descriptors can be more informative. For a property p_i associated with cation i, the composition-weighted mean is given by Equation (6):

$$\overline{p} = \sum_{i=1}^{n} x_i p_i \tag{6}$$

The corresponding composition-weighted variance is given by Equation (7):

$$\sigma_p{}^2 = \sum_{i=1}^{n} x_i (p_i - \overline{p})^2 \tag{7}$$

Such descriptors allow an ML model to distinguish between the average chemical environment and also the degree of chemical heterogeneity. For example, two- and five-cation HEOs can have identical configurational entropy but substantially different ionic-radius variance or electronegativity variance. Their structural and electronic behavior may consequently differ even though both possess the same ideal ΔS_{config} [35,104].

3.3.2. Thermodynamic Descriptors

Thermodynamic descriptors are particularly important for predicting HEO phase stability. Configurational entropy, expressed again in Equation (8), is the most familiar descriptor:

$$\Delta S_{\text{config}} = -R \sum_{i=1}^{n} x_i \ln x_i \tag{8}$$

However, recent ML studies demonstrate why entropy alone is insufficient. Sivak et al. [9] used mixing enthalpy together with bond-length distribution to map rock-salt HEO stability. Their results showed that these descriptors could successfully distinguish experimentally stabilized single-phase compositions and identify additional candidates. Dicks et al. [8] similarly used enthalpy of mixing as an enthalpic descriptor and introduced the variance of individual cation energies as an entropy-related descriptor. Their approach explicitly incorporated both enthalpic and entropic contributions rather than assuming that configurational entropy alone determines HEO stability. Other potentially useful thermodynamic descriptors include formation energy, decomposition energy, energy above the convex hull, chemical potential, and defect formation energies. The selection of descriptors should therefore reflect the specific prediction problem. For example, descriptors relevant to phase stability may not be identical to those needed for catalytic activity or electronic conductivity.

3.3.3. Structural Descriptors

Structural descriptors describe the geometric environment created by multination incorporation. Important variables include lattice parameters, coordination number, bond lengths, bond angles, radius mismatch, structural symmetry, density, packing characteristics, and local-environment fingerprints. Among these, bond-length distributions are particularly relevant to HEOs because chemical disorder produces a distribution of local bond environments even when the average crystal structure remains well defined. Sivak et al. [9] explicitly demonstrated the utility of bond-length distribution in combination with mixing enthalpy for distinguishing single-phase rock-salt HEOs.

The importance of structural descriptors also explains why purely composition-based ML models may eventually reach a performance limit. The same elemental composition can potentially produce different local configurations or phases, meaning that a model that receives only elemental fractions cannot distinguish all physically relevant states. Structure-related approaches therefore represent a progressive evolution from composition-based ML models to models that integrate composition with structural information and, ultimately, to atomistic approaches capable of capturing local chemical environments and atomic-scale interactions [105–107].

3.3.4. Electronic Descriptors

Electronic descriptors become important when the target property involves optical, catalytic, magnetic, or electronic behavior. Potential descriptors include band gap, work function, electron affinity, d-electron count, Fermi energy, ionization potential, orbital occupancy, and density-of-states-derived quantities. For HEOs, the challenge is that electronic behavior may not be represented adequately by the average electronic properties of the constituent elements. Local coordination, oxidation state, cation oxygen bonding, oxygen vacancies, and lattice distortion can all modify the electronic structure [108].

Graph-based ML approaches provide one possible solution because they can learn structural and chemical relationships directly from atomic environments. The crystal graph convolutional neural network (CGCNN) introduced by Xie and Grossman [109] represents crystalline materials through atoms and their connections rather than requiring a manually constructed fixed descriptor vector. The model demonstrated accurate prediction of multiple materials properties and provided access to contributions from local chemical environments. Similarly, MEGNet [110] treats materials as graph networks and demonstrates accurate prediction of properties such as formation energy, band gap, and elastic properties across large crystal datasets. These approaches are particularly promising for HEOs because local chemical environments are themselves a major source of HEO functional behavior [109,110].

3.3.5. Defect Descriptors

Defect chemistry constitutes an underdeveloped but potentially critical descriptor space for HEO machine learning. Oxygen-vacancy formation energy, vacancy concentration, vacancy charge state, cation-vacancy formation energy, oxygen migration barriers, and oxygen chemical potential are potential inputs for models aimed at predicting redox, catalytic, ionic-conductivity, or electronic properties. The importance of defect descriptors arises from the fact that HEOs can accommodate multiple oxidation states and charge compensation mechanisms. Consequently, a composition may determine not only which cations are present but also the likelihood of oxygen non-stoichiometry and the resulting electronic structure [111].

3.3.6. Synthesis Descriptors

For HEOs, descriptor selection should follow the physical hierarchy governing their behavior rather than being determined solely by algorithmic convenience. Table 1 summarizes the descriptor hierarchy, while Table 2 compares ML algorithms and their HEO-specific limitations. Composition-based descriptors capture elemental identity, cation fractions, configurational entropy, ionic-size mismatch, electronegativity, and oxidation-state distributions. Structure-aware descriptors add lattice parameters, coordination environments, bond-length distributions, and crystallographic information, while atomistic descriptors can represent local chemical environments, short-range correlations, and defect configurations. Processing and reaction-condition descriptors provide an additional layer because phase formation and functional performance may depend strongly on synthesis and operating conditions. This hierarchy links composition to structure, defects, electronic properties, and performance and provides a basis for developing more transferable HEO ML models [112].

A distinctive characteristic of HEOs is that the final material is not necessarily determined solely by nominal composition. Synthesis temperature, reaction time, heating rate, atmosphere, pressure, precursor chemistry, solvent, pH, and cooling rate can affect phase constitution, crystallinity, defect concentration, oxidation state, particle size, and local chemical ordering. Consequently, synthesis conditions should be incorporated into ML datasets whenever the objective is to predict synthesizability or experimentally accessible phase formation.

This issue is demonstrated particularly well by recent HEO studies. Dicks et al. [8] explicitly framed their work around predicting synthesizability, while a recent study [94] combined ML classification with experimental synthesis and structural verification of a newly predicted composition. The fluorite study is especially instructive because its model achieved weighted F_1-scores of 77% in nested stratified cross-validation and 70% during external validation before experimentally synthesizing and confirming the predicted $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$. Thus, synthesis variables should not be regarded merely as metadata. For HEO discovery, they can be treated as design variables [94].

Table 1. Hierarchical descriptors for data-driven HEO design.

Descriptor Level	Representative Descriptors	Physical Meaning	Typical Source	Typical ML Target	References
Composition	Cation fractions; S_{config}; oxygen stoichiometry	Chemical complexity and composition	Literature/experiment	Phase/structure	[53,103,113,114]
Elemental	Ionic radius; electronegativity; valence	Cation chemistry	Materials databases	Structure/stability	[53,96,115]

Table 1. *Cont.*

Descriptor Level	Representative Descriptors	Physical Meaning	Typical Source	Typical ML Target	References
Thermodynamic	ΔH_{mix}; formation energy; energy above hull	Phase stability	DFT/high throughput	Stability	[9,35]
Structural	Lattice parameter; bond length; coordination	Lattice distortion and structure	XRD/DFT/MLIP	Structure	[9,22,77]
Local/atomistic	Neighbor identity; local coordination	Chemical heterogeneity	DFT/GNN/MLIP	Defects/properties	[59,84,89]
Defect	E_{VO}; vacancy concentration	Defect stability and chemistry	DFT/EPR/XPS	Catalytic properties	[10,116,117]
Processing	Temperature; atmosphere; precursor; time	Experimental accessibility	Literature/experiment	Synthesizability	[8,35,94]
Reaction	Light; pH; pressure; feed composition	Operating environment	Experiment	Performance	[118–120]

Table 2. Machine learning algorithms, representative uses, and principal HEO-specific limitations.

Algorithm	Principal Advantage	Potential HEO Use	HEO-Specific Limitation	References
Linear regression	Simple and interpretable	Baseline composition-property relationships	Simple baseline but unable to capture strong nonlinear interactions or configurational degeneracy.	[121–123]
Random forest	Robust nonlinear modeling	Phase/property classification	Robust on small tabular datasets but extrapolation outside sampled chemistry can be weak.	[124]
SVM	Effective for small datasets	Phase classification	Effective in limited/high-dimensional data but sensitive to kernel and feature scaling choices.	[104,125]
Gradient boosting/XGBoost	Strong tabular-data performance	Phase and property prediction	Strong tabular performance; random splits can still hide chemical leakage.	[7,94,98,126]

Table 2. *Cont.*

Algorithm	Principal Advantage	Potential HEO Use	HEO-Specific Limitation	References
ANN/deep learning	Highly nonlinear relationships	Large datasets	Requires larger datasets and careful regularization; interpretability and calibration are important.	[113,114,127,128]
GNN	Direct structural representation	Structure-property prediction	Represents local structure directly, but one nominal HEO composition may require an ensemble of graphs.	[10,37–39,91,92,109, 110,129]
MLIP	Atomic-level energy/force prediction	Structure relaxation and screening	Enables atomistic relaxation and sampling, but reliability depends on training coverage of charge, defects, magnetism, and local environments.	[8,9,93]

3.3.7. Graph Neural Networks and Atomistic ML

Graph neural networks represent materials as networks of atoms and their interactions. This is conceptually attractive for HEOs because the properties of compositionally disordered structures depend strongly on local neighborhoods. For instance, CGCNN [109], MEGNet [110], ALIGNN [129], M3GNet [39], CHGNet [38], and equivariant structure [90–92] demonstrate progressively richer ways of encoding atomic identity, connectivity, geometry, charge-sensitive environments, energies, and forces. For HEOs, however, one graph is rarely a complete representation of one composition because different cation arrangements, site occupancies, vacancies, and magnetic/oxidation states can coexist at similar bulk stoichiometry [57,59,61,66]. A disorder graph workflow should therefore sample or marginalize over a physically representative configurational ensemble rather than assuming that a single arbitrary supercell uniquely defines the material. ML interatomic potentials provide an even more atomistically detailed route. The HEO studies of Sivak et al. [9] and Dicks et al. [8] show how MLIPs can approximate complex potential energy surfaces and expand composition/structure screening beyond what brute-force DFT can feasibly cover [8,9]. Their use should nevertheless be accompanied by tests for extrapolation to evaluate local chemistry [93], oxygen-vacancy states [102], redox configurations [38], and magnetic order [25,38,93,102].

3.4. Unsupervised Learning and Composition-Space Mapping

Supervised ML requires a defined target variable, whereas unsupervised learning seeks to identify patterns within the data without assigning a predefined target. This makes unsupervised approaches particularly useful for visualizing and organizing the enormous HEO chemical space. For instance, principal component analysis (PCA) can reduce a large descriptor set to a smaller number of orthogonal components representing the dominant sources of variance. In HEO studies, PCA can reveal whether compositions cluster primarily vary according to elemental chemistry, structural family, entropy, ionic-size mismatch, or

other descriptors. K-means and hierarchical clustering can then be used to identify groups of chemically similar compositions. Such clustering may help reveal families of HEOs that share common descriptors or structural characteristics even when they belong to different nominal compositions.

Nonlinear dimensionality reduction approaches such as t-SNE and UMAP are particularly useful when the relationships among compositions cannot be represented adequately in two or three linear dimensions. However, these methods should primarily be interpreted as visualization tools rather than direct evidence of physical similarity. Apparent clusters can depend on preprocessing choices, distance metrics, hyperparameters, and feature selection. Unsupervised learning can map a high dimensional descriptor space into a lower dimensional chemical landscape, enabling visualization of compositional similarities and identification of sparsely explored regions for candidate selection. This approach can complement supervised models by revealing poorly represented regions of the HEO chemical space, where model predictions are generally associated with greater uncertainty [130].

3.5. Model Evaluation and Validation

Model accuracy needs to be interpreted carefully in HEO ML because a high statistical score does not necessarily imply successful prediction of genuinely unexplored compositions [25,28]. The conventional division into training, validation, and test datasets remains essential, but discovery-oriented studies should additionally ask whether the test samples are chemically separated from the training distribution. Random splitting is particularly permissive for HEO datasets because neighboring compositions can share most of their elements and descriptor values, causing models to be evaluated primarily on interpolation rather than chemical extrapolation [25,28,97]. For regression, mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2) remain useful descriptive metrics, but they ought to be complemented by task-relevant measures when the scientific objective is candidate discovery, ranking, or stability classification [25,26,131]. These metrics are defined in Equations (9)–(11).

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{11}$$

Here, N is the number of evaluated samples, y_i and $\hat{y}_i$ are the observed and predicted values, respectively, and $\overline{y}$ is the mean observed value.

For classification, accuracy alone is often insufficient, particularly when phase classes are imbalanced [131]. A recent study provides a useful example: nested stratified cross-validation produced a weighted F1-score of 0.77, whereas external validation decreased the score to 0.70 before a model-selected $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$ composition was synthesized and verified by XRD and TEM/SAED [94]. This type of external validation is more informative for discovery than an isolated random-split score. More demanding tests include leave-one-element-out leave-one-composition-family-out, structure-family holdout, chronological literature splits, and distance-based splitting in descriptor or learned representation space [28].

3.6. Explainable and Interpretable ML

High predictive accuracy alone is insufficient for materials discovery. A black box model that predicts a stable HEO without revealing why the composition is favorable

provides limited chemical understanding. Explainable ML (XAI) addresses this limitation by connecting model predictions with the descriptors that drive them.

Feature importance analysis can provide a first indication of which descriptors contribute most strongly to a model. However, global feature importance does not necessarily reveal how a particular descriptor affects individual predictions. SHAP (SHapley Additive exPlanations) provides a more detailed framework by assigning a contribution value to each feature for an individual prediction. Lundberg and Lee developed SHAP as a unified framework for interpreting complex ML models [132]. For HEOs, SHAP can be particularly valuable because it can translate a prediction into a chemically interpretable statement. For example, rather than reporting only that a composition has a high probability of forming a single phase, SHAP analysis can indicate whether the prediction is primarily associated with electronegativity, electron affinity, ionic-size characteristics, thermodynamic descriptors, or other variables. This capability has already been demonstrated in recent HEO research. The study by Santos et al. [94] used SHAP to identify the physical relevance of model predictors and reported strong influences from electron affinity and electronegativity across multiple descriptor scales.

Other interpretability tools, including partial-dependence plots and accumulated-local-effect plots, can further reveal whether a descriptor has a monotonic, nonlinear, or threshold-like relationship with the predicted property. The scientific importance of explainability is therefore greater than simply making an ML model easier to understand. In HEO research, XAI can generate new chemical hypotheses. If multiple independently trained models consistently identify ionic-radius variance, electronegativity distribution, or mixing enthalpy as important predictors, these descriptors can provide experimentally testable hypotheses about the mechanisms controlling phase stability. Consequently, explainable ML can extend materials prediction beyond statistical performance by enabling interpretation of model outputs, identification of chemically meaningful relationships, formulation of testable hypotheses, and subsequent experimental verification, thereby transforming ML from a predictive tool into a mechanism-oriented approach for HEO discovery [133,134].

3.7. From Predictive ML to Data-Driven HEO Discovery

Figure 8 summarizes a closed-loop HEO discovery workflow linking model training, candidate ranking, synthesis, characterization, and dataset expansion. Early materials-informatics applications largely focused on predicting properties of known materials. For HEOs, however, the more substantial opportunity is to use ML to navigate chemical space that has not yet been experimentally explored. A modern HEO discovery workflow can therefore be represented as follows:

The experimental result then becomes new training information, creating an iterative experiment–model–experiment feedback loop. The recent HEO literature already provides early examples of this paradigm. Sivak et al. [9] used MLIP-enabled screening to identify additional rock-salt candidates beyond experimentally known compositions. Dicks et al. [8] expanded the approach to fourteen elements and multiple crystal structures while explicitly considering synthesizability. The fluorite study progressed further by predicting a composition and experimentally synthesizing and structurally validating the predicted HEO. Meanwhile, automated literature mining has demonstrated a potential route for continuously expanding HEO datasets from previously inaccessible unstructured information. Thus, the most important role of ML in HEO research is not simply to provide a faster alternative to conventional regression. Its greater potential lies in establishing a closed-loop discovery framework in which existing knowledge is systematically encoded, unexplored chemical space is computationally prioritized, predictions are experimentally tested, and the resulting observations are fed back into the model [135,136].

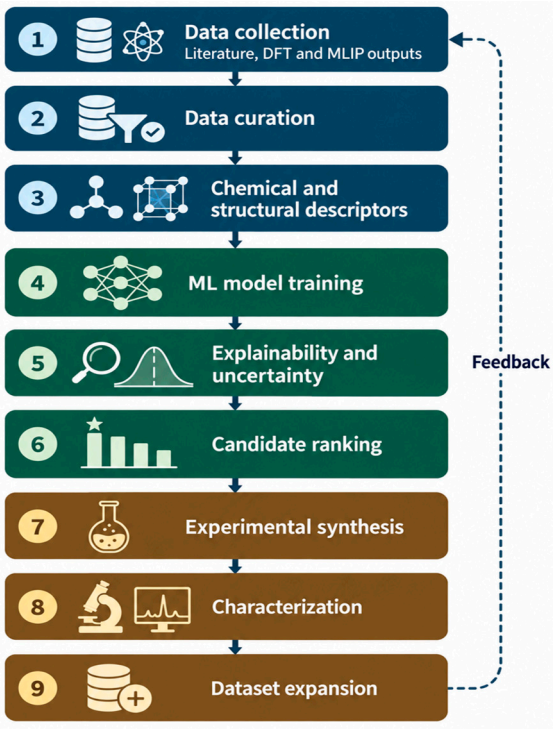


Figure 8. Closed-loop workflow for data-driven HEO discovery.

This perspective also exposes an important limitation. HEO datasets remain relatively small compared with the enormous number of theoretically possible compositions. Consequently, model uncertainty, dataset bias, descriptor redundancy, chemical extrapolation, and experimental inconsistencies remain major barriers. The future development of HEO informatics will therefore require not only increasingly sophisticated algorithms but also larger and better-curated datasets, physically meaningful descriptors, uncertainty-aware prediction, standardized synthesis metadata, and systematic experimental validation [137,138].

4. Machine Learning-Guided HEO Prediction and Discovery

The fundamental complexity of high-entropy oxides (HEOs) arises from the simultaneous variation in elemental composition, cation ratio, oxidation state, crystal structure, local coordination, defect chemistry, and synthesis conditions. As discussed in the preceding sections, conventional descriptor-based approaches can provide useful guidelines for identifying potentially stable compositions, but they become increasingly inadequate as the number of chemical and structural variables increases. In this context, machine learning (ML) offers an alternative strategy in which experimentally and computationally generated data can be used to establish nonlinear relationships between composition, structure, processing conditions, and properties. Importantly, the objective of ML-assisted HEO is to enable the systematic exploration of previously inaccessible regions of the HEO design space [24,139].

Recent studies demonstrate that ML can already contribute to several stages of HEO discovery, including crystal-structure classification, stability and synthesizability screening, prediction of structural parameters, oxygen-vacancy engineering, and composition optimization. For example, ML interatomic potentials have been used to rapidly map rock-salt HEO composition space, while supervised ML models have demonstrated high accuracy in crystal-structure classification [7,9]. More recently, ML has been coupled with DFT and experimental validation to identify oxygen-vacancy-rich high-entropy perovskites and to discover previously unexplored HEO compositions. The development of these approaches

represents a transition from composition-by-composition screening toward data-driven navigation of multidimensional chemical space. The following sections critically examine how this transition is occurring and where important limitations remain [140].

4.1. ML Prediction of Phase Stability and Synthesizability

One of the most important challenges in HEO discovery is determining whether a selected multication composition can form a stable or metastable single-phase structure rather than separating into multiple phases. Conventional thermodynamic descriptors, including configurational entropy, mixing enthalpy, ionic-size mismatch, and related empirical parameters, provide valuable first-order guidance but do not completely capture the highly heterogeneous local environments of HEOs. ML provides an opportunity to combine multiple descriptors and, in some cases, to represent the atomic interactions directly. A notable advancement was reported by Sivak et al. [9], who developed a self-consistent computational–experimental framework based on a machine learning interatomic potential for exploring single-phase rock-salt HEOs. Rather than relying exclusively on configurational entropy, their approach used bond-length distribution and mixing enthalpy as descriptors of the chemically disordered system. The ML interatomic potential enabled rapid exploration of a large number of compositions at a computational cost substantially lower than performing direct DFT calculations for every disordered configuration. Importantly, the framework correctly resolved the single-phase stability of experimentally stabilized rock-salt compositions and identified additional compositions as candidates for experimental investigation [9].

This work is significant because it illustrates an important conceptual shift. Phase stability in an HEO cannot necessarily be inferred from configurational entropy alone. Local structural accommodation, reflected through the distribution of bond lengths, provides complementary information about whether the multication lattice can tolerate chemical disorder. Consequently, ML can integrate thermodynamic and local structural descriptors into a unified stability map. Dicks et al. [8]. subsequently extended this concept toward prediction of HEO synthesizability. Their methodology employs ML interatomic potentials to construct and relax large random unit cells across different elemental combinations and crystal structures. Candidate compositions are evaluated using descriptors based on the variance of individual cation energies and the enthalpy of mixing. The approach was tested on tetravalent HEOs and subsequently extended to 14 elements and three crystal structures. Importantly, the method identified the known stable four-component HEO in the α-PbO_2 structure and predicted several previously unexplored five-component candidates.

These studies demonstrate that ML interatomic potentials can address a key bottleneck in HEO discovery: the computational cost of explicitly sampling chemical disorder [8,9]. Nevertheless, computational stability and experimental synthesizability must remain distinct targets. A composition predicted to be favorable under a particular structural model may still fail experimentally because precursor chemistry, kinetic pathways, oxygen chemical potential, temperature, atmosphere, cooling rate, and synthesis history alter the accessible free-energy landscape [5,35,94,141]. MLIP predictions are therefore most defensible as candidate-ranking tools whose uncertainty and domain of validity are reported, followed by DFT checks where appropriate and independent experimental synthesis. Literature-mining approaches provide a complementary route by transforming composition-phase-synthesis information into machine-readable records, but these datasets inherit publication bias, inconsistent phase labels, and missing negative outcomes [33,34,98,142].

4.2. ML Prediction of Crystal-Structure Selection

The ability to predict the likely crystal structure from composition is particularly valuable for HEOs because the same multication chemistry can potentially accommodate different structural families. Rock-salt, fluorite, spinel, perovskite, pyrochlore, and other structures exhibit substantially different cation coordination environments and therefore different thermodynamic and functional properties [94].

Liu et al. [7] directly addressed this problem by developing ML models for HEO crystal-structure prediction. Three conventional ML algorithms were evaluated using multiple performance metrics, with the XGBoost classifier showing particularly strong predictive performance, reaching reported accuracy and F1-scores of 0.977 and 0.975, respectively. Importantly, the study did not simply demonstrate predictive accuracy; its SHAP analysis provided insight into the chemical descriptors controlling structural classification. The anion-to-cation radius ratio was identified as the most influential descriptor, followed by differences in Pauling and Mulliken electronegativities. This finding is particularly important for HEO design because it illustrates how explainable ML can convert statistical correlations into chemically interpretable design rules. Rather than treating the ML model as a black box, feature-importance analysis can indicate which combinations of geometric and electronic descriptors are associated with specific structural families.

A more recent study by Santos et al. [94] further expanded the phase classification problem. Their database included multiple HEO structural categories, and ML was used to distinguish between different phase outcomes. Their work demonstrates the possibility of moving from binary questions such as single-phase versus multiphase toward multiclass prediction of structure selection. The significance of this development extends beyond phase classification because an ML model capable of linking composition with the probability of forming a particular crystal structure can also facilitate inverse structure-guided design, in which the desired structural family is specified first and ML is then used to identify compositional and chemical descriptors associated with that structure, followed by screening of candidate compositions. However, structure prediction should not be interpreted as a purely compositional problem. Experimental processing conditions can alter phase formation, particularly in kinetically controlled systems. Therefore, future structure-prediction models should incorporate composition, processing and synthesis history, rather than relying exclusively on elemental descriptors.

4.3. Prediction of Lattice Distortion and Local Structural Environments

The average crystallographic structure of an HEO provides only a partial description of its actual atomic environment. Because multiple cations with different radii, valence states, and bonding characteristics occupy the same crystallographic sublattice, individual metal–oxygen bonds can exhibit substantial local variation. Consequently, the relationship between composition and structure must be considered at both the average and local scales.

The DFT-ML study by Lin et al. [143] provides an early example of this approach for the spinel $(Co,Cr,Fe,Mn,Ni)_3O_4$ system. First-principles calculations were combined with ML to predict lattice constants and formation energies for stoichiometric and nonstoichiometric structures. The study demonstrated that lattice constants depend on composition and that cation-specific site occupation can substantially influence the calculated energetics.

More recently, Sivak et al. [9] demonstrated the usefulness of bond-length distributions for mapping rock-salt HEO stability. Instead of describing disorder solely through a scalar configurational-entropy value, the distribution of local bond lengths provides information about how effectively the lattice accommodates different cations. This distinction is important because HEOs with similar nominal configurational entropies can exhibit substantially different local structural strain and chemical environments. Accordingly, ML models can

incorporate descriptors that capture average lattice parameters, cation oxygen bond-length distributions, local coordination environments, lattice strain, structural heterogeneity, and cation-specific distortions, thereby providing a more detailed representation of the structural complexity underlying HEO phase stability and functional behavior. Future models should ideally move beyond composition-averaged descriptors and explicitly represent local atomic environments. Graph-based ML models and ML interatomic potentials are particularly promising in this regard because they can encode relationships among neighboring atoms rather than treating the material simply as a vector of elemental averages [39,91,106].

4.4. ML-Guided Defect and Oxygen-Vacancy Engineering

Defects represent one of the most important links between HEO composition and functional performance. In oxide materials, oxygen vacancies can modify oxidation states, local coordination, electronic structure, adsorption behavior, and catalytic reaction pathways. However, predicting oxygen-vacancy formation in HEOs is considerably more complicated than in conventional oxides because multiple chemically distinct cations and local environments coexist within the same lattice. A significant development was reported by Tukur et al. [10], who introduced OxiGraphX, a graph neural network designed to predict oxygen-vacancy formation energies in high-entropy perovskite oxides (HEPOs). Their framework explicitly combined ML, DFT and experimental validation. The model was trained to establish relationships among composition, atomic structure, and local chemical environments and was subsequently used to screen compositions for favorable oxygen-vacancy characteristics.

The importance of this approach lies in its potential to shift defect engineering from a predominantly empirical strategy toward composition-guided defect design. Rather than synthesizing numerous HEO compositions and subsequently evaluating oxygen-vacancy concentrations, ML can identify compositions with favorable vacancy-formation energetics before experimental synthesis [10]. However, oxygen-vacancy formation energy and experimentally realized vacancy concentration are not interchangeable quantities: vacancy populations additionally depend on temperature, oxygen partial pressure, charge compensation, defect interactions, synthesis atmosphere, local cation configuration, and kinetic history [84,116,117]. The OxiGraphX study illustrates an appropriate validation chain by relating ML-predicted vacancy formation energies to DFT, experimentally determined oxygen-vacancy content, and OER response; the reported mean absolute error for average OVFE prediction was approximately 0.023 eV, while experimentally tested compositions showed an ordered OER response consistent with the predicted vacancy trend [10]. Experimental techniques such as XPS, EPR, XAS, Raman spectroscopy, isotope exchange, and thermogravimetry remain essential for determining whether the predicted defect tendency is actually realized [117,144].

4.5. ML Prediction of Electronic and Optoelectronic Properties

HEO electronic properties are strongly influenced by multication disorder, mixed valence, local coordination, lattice distortion, and defects. These interactions make direct prediction of properties such as band gap, band-edge positions, density of states, and work function particularly challenging. The earliest DFT-ML workflow for a photocatalytically relevant HEO, the spinel $(Co,Cr,Fe,Mn,Ni)_3O_4$ study of Lin et al. [143] discussed in Section 4.3, showed that ML could reproduce composition-dependent formation energies and lattice constants while substantially reducing the cost of direct first-principles calculations for every configuration. However, that study predicted structural and energetic quantities rather than band gaps or band-edge positions, so it provides a methodological basis for, rather than a demonstration of, ML prediction of HEO electronic structure.

For optical properties, Hoyos-Sanchez et al. [108] recently developed an ML framework for automated determination of HEO band-gap energies from UV-vis spectroscopy data. Their workflow included preprocessing, data augmentation and model optimization. Bagging, Extra Trees and random forest models achieved reported MAEs of approximately 0.26–0.28 eV and R^2 values of approximately 0.73–0.74. This study is valuable, but its scientific interpretation should be precise. The model primarily accelerates band-gap extraction from spectroscopic data, rather than constituting a universal composition-to-band-gap predictive model. This distinction is important when assessing the maturity of ML-based electronic-property prediction in HEOs.

Future models could integrate composition, local coordination environments, oxidation states, and defect-related descriptors to establish more comprehensive relationships with the electronic structure of HEOs, thereby improving the prediction and interpretation of composition-dependent electronic properties. Such models could ultimately predict band gaps, band-edge positions, density-of-states characteristics, work functions, carrier-related descriptors, and adsorption-relevant electronic states. However, these capabilities remain less mature for HEOs than phase classification and stability prediction. Therefore, experimental and first-principles validation remain essential.

4.6. Data-Driven Optimization of Multicomponent HEO Composition

Most early HEO studies focused on equimolar systems because equal elemental fractions provide a straightforward route to maximizing ideal configurational entropy. However, maximum configurational entropy is not necessarily equivalent to optimum functional performance or optimum phase stability. Consequently, ML provides an opportunity to move from the conventional equimolar design concept toward compositionally optimized, non-equimolar HEOs.

Instead of restricting the composition to $(A_{0.2}B_{0.2}C_{0.2}D_{0.2}E_{0.2})$ Oδ, the design space can be expressed more generally as (AxByCzDmEn)Oδ, with x + y + z + m + n = 1. Here, the ML model can simultaneously consider elemental identity, concentration, oxidation state, oxygen non-stoichiometry and structural family. The importance of this approach is demonstrated by Sivak et al. [9], whose MLIP-guided exploration of rock-salt HEOs correctly resolved the single-phase stability of all experimentally known equimolar compositions and led to a previously unreported non-equimolar, Ca-containing rock-salt HEO. Similarly, Santos et al. [94] used ML to identify a previously unexplored single-phase fluorite HEO, $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$, which was subsequently synthesized and characterized by XRD and TEM/SAED. The study also used SHAP analysis to identify electron affinity and electronegativity-related descriptors as important factors in phase formation [94].

4.7. Toward Multi-Objective Optimization of HEO Performance

The ultimate objective of HEO design is rarely the maximization of a single material property. Practical materials must simultaneously satisfy several requirements. Figure 9 illustrates the simultaneous optimization of phase stability, activity, operational stability, defect chemistry, cost, and sustainability. For example, an electrocatalyst should combine high intrinsic activity with long-term stability, while an energy-storage material may require high capacity, structural stability, favorable kinetics and low cost. However, fully developed multi-objective optimization of HEOs remains an emerging rather than mature field. Current HEO ML studies have largely concentrated on phase stability, structure, defect descriptors, or individual target properties. Therefore, multi-objective optimization should be presented as the next logical stage rather than as an already established general methodology [140]. A future framework could generate a Pareto-optimal set of compositions rather than a single "best" HEO: Such an approach would be particularly powerful for environmental and energy

applications, where catalytic activity, durability, elemental criticality and environmental impact must be considered simultaneously.

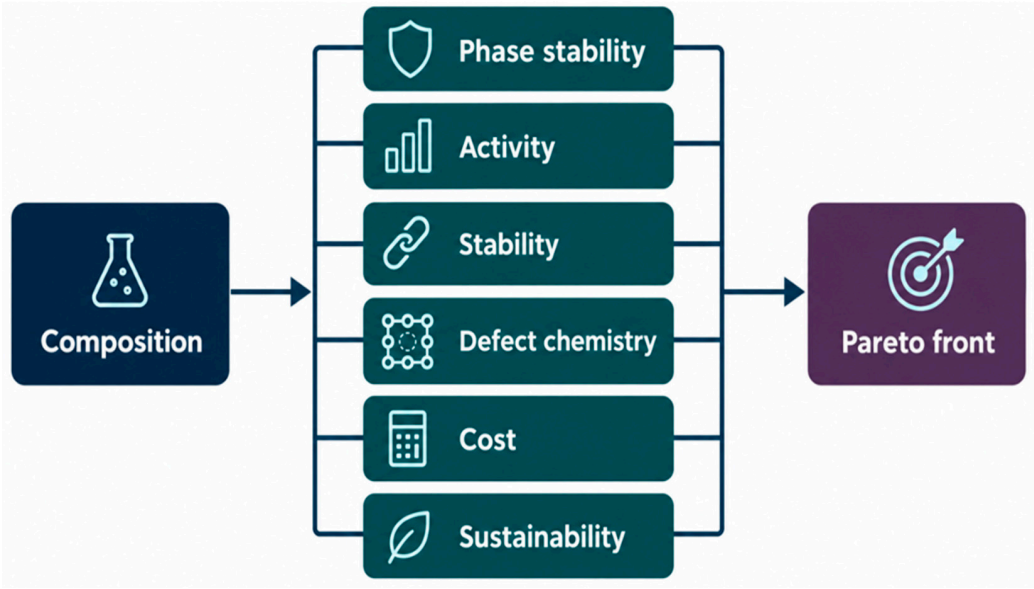


Figure 9. Conceptual framework for multi-objective optimization of HEO composition.

4.8. Coupling ML, DFT, and ML Interatomic Potentials for Accelerated Screening

The combination of ML with DFT provides one of the most practical routes toward reliable HEO screening. DFT offers relatively high-fidelity information about energetics and electronic structure but is computationally expensive, especially for chemically disordered supercells containing many possible cation arrangements. ML can learn from a limited number of high-quality calculations and subsequently predict properties for much larger composition spaces.

An early demonstration was provided by Lin et al. [143], who used first-principles calculations to generate data for spinel HEOs and subsequently trained ML models to predict lattice constants and formation energies. The authors demonstrated that the combined approach could reduce computational requirements compared with performing first-principles calculations for every configuration. The development of ML interatomic potentials represents a further step. Sivak et al. [9] used an ML interatomic potential to rapidly map rock-salt HEO composition space, while Dicks et al. [8] applied ML interatomic potentials to screen substantially larger elemental and structural spaces and predict candidate HEOs.

For defect engineering, Tukur et al. [10] provide another important example in which DFT-generated oxygen-vacancy formation energies were used to train a graph neural network, followed by computational screening and experimental validation. Thus, the emerging computational framework involves using DFT calculations to generate high-quality training data, followed by ML or ML interatomic potential (MLIP) development for rapid prediction and large-scale screening, candidate selection based on the predicted properties or stability, and subsequent DFT verification and experimental validation. The key advantage is therefore not that ML replaces DFT but that it changes how DFT resources are allocated. DFT becomes a source of high-quality reference information, whereas ML enables exploration of regions that would be computationally inaccessible through brute-force calculations [10].

4.9. Active Learning and Bayesian Optimization for Closed-Loop HEO Discovery

Conventional supervised ML generally follows a passive learning paradigm: a fixed dataset is collected, a model is trained, and predictions are generated for new compositions. However, this approach is inefficient when the design space is extremely large and data acquisition is expensive. Active learning and Bayesian optimization offer a fundamentally different strategy by allowing the model to determine which candidate should be evaluated next. The basic active-learning cycle involves training an initial model using an existing dataset, estimating prediction uncertainty, selecting the most informative or promising candidates, evaluating them through new calculations or experiments, and incorporating the newly acquired data to iteratively update and improve the model [145]. This process is particularly suitable for HEOs because only a very small fraction of the theoretically possible compositions can realistically be synthesized and characterized.

A recent study by Brorsson et al. [29] demonstrated this strategy for high-entropy oxygen carriers. Their active-learning framework integrates predictive modeling, uncertainty estimation and iterative sampling to explore the embedded composition space. Rather than evaluating candidates uniformly, the algorithm focuses subsequent searches on promising regions and was reported to accelerate the discovery of high-entropy oxygen carriers relative to conventional trial-and-error and grid-search approaches. Bayesian optimization provides a related strategy in which an acquisition function balances exploration of poorly understood regions against exploitation of compositions predicted to have desirable properties. Bayesian optimization has been used to locate stable short-range-ordered configurations in HEOs [83] and, in Pareto-guided multi-objective form, to steer a robotic AI-chemist toward high-entropy chlorine-evolution catalysts [140]; in HEA electrocatalysis, iteratively refined computational models combined with composition-spread thin-film libraries correctly located the activity maxima of the Ag–Ir–Pd–Pt–Ru system for the oxygen reduction reaction [146].

4.10. From Virtual Screening to Inverse HEO Design

The progression of ML-assisted HEO research is gradually shifting the field from conventional property prediction toward inverse materials design. In conventional materials research, the development process generally follows a forward sequence in which the composition determines the structure, the structure governs the resulting properties, and these properties ultimately determine the material's functional performance [8].

In contrast, inverse materials design begins with a targeted property or functional performance and works backward to identify a suitable material composition. In this approach, an ML model maps the desired target to promising candidate compositions, which are subsequently evaluated in terms of their likely structural and defect configurations and finally validated through experimental synthesis. Recent HEO studies provide early examples of this transition. The ML-assisted fluorite HEO study demonstrated that a model trained on known composition-phase relationships could identify a new composition that was subsequently synthesized and experimentally confirmed as a single-phase fluorite HEO. Figure 10 summarizes the progression from prediction and screening to discovery, validation, active learning, and inverse/autonomous design. Similarly, the literature-mining/ML framework developed by Santos et al. [98] illustrates how structured composition-phase information can be transformed into predictive design rules for single-phase HEOs.

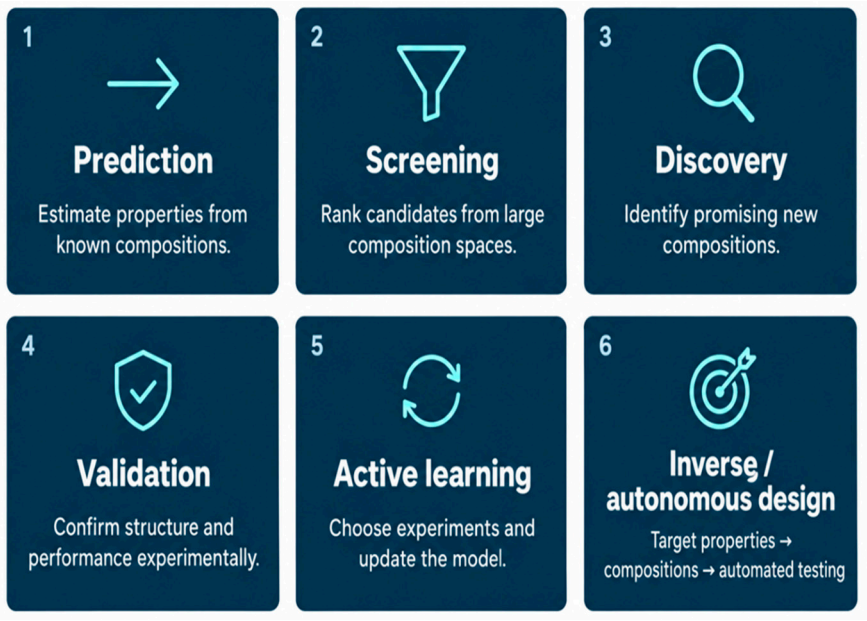


Figure 10. Progression from predictive ML to screening, discovery, active learning, and inverse/autonomous HEO design.

The MLIP studies of Sivak et al. [9] and Dicks et al. [8] further demonstrate that very large hypothetical composition spaces can be explored computationally before experimental resources are committed. Nevertheless, true inverse HEO design remains more challenging than conventional prediction. A model may identify a composition with a high predicted probability of possessing a particular property, but this does not guarantee successful synthesis or the expected performance. The inverse-design problem therefore needs to incorporate uncertainty, synthesizability, processing conditions, and experimental feedback [94].

The long-term vision is therefore to move beyond standalone ML prediction toward a closed-loop, uncertainty-aware discovery framework in which literature and DFT data are used to train ML models, guide candidate screening and active-learning decisions, and prioritize compositions for synthesis and characterization. The resulting experimental data can then be incorporated into the model to iteratively improve its predictive capability, potentially transforming HEO discovery from a predominantly trial-and-error process into a self-improving materials-discovery workflow.

5. Functional Applications and Condition Performance Optimization

HEO applications in catalysis, photocatalysis, electrochemical energy storage, and environmental remediation provide a rich experimental basis for data-driven materials design. However, the maturity of ML integration differs substantially among these areas. Phase stability, structural descriptors, defect energetics, and selected physicochemical properties are currently more amenable to ML prediction than final application-level metrics such as catalytic selectivity, photocatalytic H_2 evolution, CO_2 conversion, pollutant degradation, or battery lifetime. Accordingly, this chapter distinguishes established HEO functionality from the emerging use of ML to predict and optimize functional performance.

High-entropy oxides (HEOs) have emerged as a versatile class of multicationic materials for energy conversion, energy storage, catalysis, photocatalysis, and environmental remediation. Their functional diversity originates from the ability to incorporate multiple chemically distinct cations within a common oxide framework, thereby creating a broad range of local coordination environments, oxidation states, lattice distortions, defect config-

urations, and electronic structures. These characteristics provide opportunities for tuning catalytic and electrochemical properties beyond the compositional flexibility of conventional oxides. Reviews of HEOs have documented their application in thermal catalysis, electrocatalysis, electrochemical energy storage, and photocatalysis, while recent work has increasingly emphasized the coupling between local structure, defect chemistry, and performance [72].

The conventional development of HEO-based functional materials, however, has generally proceeded through experimental selection of a limited number of elemental combinations followed by synthesis, characterization, and performance testing. This approach samples only a very small fraction of the theoretically accessible composition space. The emergence of machine learning (ML) provides a means of connecting the established HEO application landscape with systematic materials discovery by using chemical, structural, thermodynamic, defect, and processing descriptors to identify promising compositions before experimental synthesis. Recent ML studies have already demonstrated this concept for phase selection, oxygen-vacancy engineering, compositional screening, and active learning, although direct ML prediction of application-level metrics such as photocatalytic H_2 evolution or CO_2 conversion remains comparatively less mature [29,30,147].

5.1. HEOs for Heterogeneous Catalysis

The multicationic composition of HEOs provides a large population of chemically distinct surface environments, making these materials attractive for heterogeneous catalysis. The coexistence of multiple cations can modify metal–oxygen bonding, oxygen mobility, oxidation states, defect populations, and adsorption characteristics, while the ability to maintain a single crystalline phase can allow chemically different functionalities to be incorporated within one material. A comprehensive review of HEO catalysis has highlighted their potential in thermal, electro-, and photocatalysis and associated their catalytic behavior with lattice distortion, oxygen mobility, oxygen-vacancy populations, and high elemental dispersion [49,72].

One representative example is the fluorite HEO developed by Riley et al. [148] for CO oxidation. The authors deliberately selected a multicationic composition to simultaneously address three requirements, i.e., catalytic activity, oxygen storage capacity, and thermal stability, rather than optimizing only one property. The resulting Ce-containing HEOs were synthesized by a relatively low-temperature sol–gel route and maintained a fluorite single-phase structure. The study found that first-row transition-metal cations were particularly effective in improving CO oxidation activity, whereas rare-earth cations contributed to thermal stability. This demonstrates an important principle for data-driven HEO design: different constituent elements may play different functional roles, and the optimum composition may arise from balancing several properties rather than maximizing one descriptor.

This concept is particularly relevant to ML because catalytic HEOs often exhibit property trade-offs. For example, introducing a cation that improves redox activity may simultaneously decrease thermal stability, whereas another cation may improve structural robustness at the expense of activity. Such relationships are difficult to identify systematically using a small number of experimentally synthesized compositions but can potentially be learned from sufficiently diverse datasets.

HEOs have also been investigated for electrocatalytic reactions including the hydrogen evolution reaction (HER), oxygen evolution reaction (OER), oxygen reduction reaction (ORR), and CO_2 reduction reaction (CO_2RR). Recent reviews emphasize the importance of structural diversity, oxygen vacancies, cation interactions, and electronic-structure modification in determining their electrochemical behavior.

For HER, HEOs can function either as catalysts themselves or as structurally/electronically active supports for highly active components. For example, a Pt/$(FeCoNiCrAl)_3O_4$ nanohybrid exhibited a low HER overpotential of 22 mV at 10 mA cm^{-2} and maintained activity for 60 h [149]. The authors attributed the improved behavior partly to electronic interactions between Pt and the HEO as well as improved Pt dispersion and reduced charge-transfer resistance. This example highlights a broader opportunity in data-driven design: the desirable HEO composition may depend not only on its intrinsic catalytic properties but also on its interaction with a second phase.

For OER, current evidence particularly emphasizes the roles of oxygen vacancies, lattice distortion, cation-site diversity and modified electronic structures. HEOs based on spinel, perovskite, rock-salt, and rutile structures have therefore been widely investigated as OER catalysts [150,151]. Importantly, however, the catalytic response cannot be reduced to a simple "more defects = higher activity" relationship. Defect concentration, defect location, oxidation state, surface reconstruction and reaction conditions jointly determine the accessible catalytic states. Surface-sensitive photoelectron spectroscopy further shows that the active redox response can be strongly cation-specific, reinforcing the need to distinguish bulk phase identity from the working surface state [152].

For CO_2RR, HEOs are attractive because multiple cations can provide different adsorption environments for CO_2 and its reaction intermediates. The large compositional space therefore creates the possibility of tuning both adsorption strength and reaction pathway. Similar principles apply to ORR and other redox reactions. The broad application space establishes HEOs as promising catalytic platforms, but it also reinforces the need for data-driven methods capable of identifying which combinations of elemental and structural descriptors actually control performance [120,153,154].

5.1.1. Lessons from High-Entropy Alloy Catalysis: Local Environments and Adsorption-Energy Distributions

Although this review focuses on HEOs, research on high-entropy alloy (HEA) catalysis has developed a quantitative framework for linking compositional complexity to catalytic activity that transfers directly to HEO design. Batchelor et al. [155] showed that the surface of a disordered HEA does not present a single adsorption energy but a near-continuous distribution of adsorption energies. Using DFT-calculated *OH and *O adsorption energies on a random subset of binding sites of IrPdPtRhRu, they fitted a simple regression model whose inputs were the identity of the binding-site atoms and the number of each element in the surrounding coordination shells, predicted the adsorption energies of all remaining sites, and combined the resulting distribution with a kinetic expression for activity to optimize the alloy composition for the oxygen reduction reaction (ORR). Because a minority of sites with near-optimal binding can dominate the overall rate, the activity is governed by the shape and position of the adsorption-energy distribution, which composition can shift. Pedersen et al. [154] extended this probabilistic approach to CO_2 and CO reduction on CoCuGaNiZn and AgAuCuPdPt by combining DFT with supervised ML to predict *CO and *H adsorption energies for all surface sites; compositions were then optimized to increase the likelihood of sites that bind *CO strongly and *H weakly, thereby suppressing H_2 evolution and favoring highly reduced carbon products. Selectivity was thus tuned through the statistics of local sites rather than through the construction of a specific site. The same group later coupled iteratively refined models with composition-spread thin-film libraries and showed that the refined model correctly predicted the experimental ORR activity maxima of Ag–Ir–Pd–Pt–Ru [146].

Saidi [156] used DFT together with an adsorbate- and site-specific deep learning model to compute the adsorption energies of *AH_x intermediates (A = C, N, O, and S) across

CoMoFeNiCu HEA surfaces and examined whether local variability allows HEAs to escape the scaling relations that limit conventional catalysts. The energies of *A and *AH_x remained linearly correlated whenever the two species occupied sites of identical symmetry, and a local linear dependence emerged between configuration-averaged adsorption energies irrespective of site symmetry. These weaker, "local" scaling relations were attributed to the nearsightedness of electronic interactions and to the narrow distribution of adsorption energies around the mean-field value, and they were sufficient to prevent independent optimization of multistep reactions. Compositional complexity alone therefore does not guarantee that scaling relations are broken.

Three lessons from HEA catalysis are directly relevant to HEO design. First, the appropriate ML target is the distribution of adsorption energies over the accessible local environments rather than a single composition-averaged value; models that map local environments to adsorption energies are therefore better suited to catalysis than direct composition-to-activity regression. Second, element-specific roles can be resolved: the identity of the binding atom sets the mean adsorption energy, while the surrounding atoms shift it, and this hierarchy has now been observed in oxides. Svane and Rossmeisl [157] transferred the HEA framework to rutile (110) HEO surfaces for the oxygen evolution reaction (OER), fitting a model to DFT reaction energies for a limited number of sites to predict reaction energies for all local environments; both the overpotential and the preferred pathway, on coordinatively unsaturated sites or via proton transfer to a bridging oxygen, depended on the local atomic environment, and an explicit surface model was used to identify the most active composition. Baek et al. [158] found by DFT that the adsorption energies of OER intermediates on spinel HEOs show wider distributions than on lower-entropy oxides because of mixing-induced strain in the active metal–oxygen bonds and confirmed experimentally that the spinel HEO outperformed lower-entropy counterparts. Song et al. [159] calculated hydrogen adsorption energies at 2000 surface O sites of rock-salt $Mg_{0.25}Ni_{0.25}Cu_{0.25}Zn_{0.25}O$ and showed that they are governed mainly by the metal cation directly beneath the O site; among several graph neural networks, DimeNet++ predicted these energies satisfactorily for this and slightly varied compositions. Mints et al. [119] further compared experimental OER activity across an HEO composition space with such theoretical models. Third, because local scaling relations may persist [156], HEO catalyst design should aim to increase the fraction of near-optimal sites and should exploit oxide-specific degrees of freedom that alloys lack, including lattice-oxygen participation, oxygen vacancies, variable cation oxidation states, and surface reconstruction.

5.1.2. Implications of Local Disorder for Reaction Mechanisms, Barriers, and Selectivity

Local chemical disorder affects not only adsorption thermodynamics but also reaction kinetics and selectivity. Transition-state energies on metal and oxide surfaces often scale with the reaction energies of the corresponding elementary steps through Brønsted–Evans–Polanyi (BEP) relations [160,161]. A distribution of adsorption energies therefore translates into a distribution of activation barriers, and because rates depend exponentially on barriers, the measured activity is dominated by the tail of the most active local configurations rather than by the average site. Oxygen-vacancy formation energies, which control Mars–van Krevelen pathways in oxidation catalysis and lattice-oxygen mechanisms in the OER, are themselves strongly dependent on the local cation environment. Chae et al. [84] showed by first-principles calculations that vacancy formation energies in entropy-stabilized oxides vary widely with the local composition and structural distortion around the vacancy, Potter et al. [116] developed a statistical description of oxygen vacancies in distorted high-entropy perovskites, and Tukur et al. [10] used a graph neural network to predict local

vacancy formation energies in perovskite HEOs, with predicted trends that matched the measured vacancy content and OER activity of the synthesized compositions.

Selectivity follows the same logic. When competing pathways depend on different intermediates, such as *CO versus *H in CO_2 reduction [154] or the conventional versus bridging-oxygen OER pathways on rutile HEOs [157], composition changes the relative populations of sites that favor each pathway. For HEO machine learning, this implies a hierarchical target. ML models or MLIPs should first predict site-resolved adsorption energies, vacancy formation energies, and, where possible, barriers over a representative ensemble of local environments; these distributions should then be coupled with mean-field or microkinetic models to estimate activity and selectivity, and only then should the predictions be compared with the experiment. Universal and equivariant MLIPs [38,39,91] make barrier sampling over many configurations increasingly tractable, but their reliability for transition states and for magnetic, mixed-valence oxide surfaces must be checked against DFT [93]. Finally, oxide surfaces are dynamic: segregation, vacancy-driven transformation, and reconstruction under reaction conditions [69,152] can change the local-environment distribution during operation. The distribution that matters for catalysis is that of the working surface, which should ideally be constrained by operando characterization [86,162].

A complementary process-level example is plasma-assisted methane valorisation. Wali et al. reviewed how energetic-electron activation, radical chemistry, reactor configuration, and operating conditions jointly influence conversion, product distribution, and energy efficiency [163]. This broader literature supports including reactor and operating-condition descriptors when modeling selectivity. Its non-equilibrium plasma mechanisms should be distinguished from thermal, electrochemical, and photocatalytic HEO mechanisms, and the review is used here as methodological context rather than evidence of ML-guided HEO discovery.

5.2. HEOs for Photocatalysis

HEO photocatalysis is a comparatively young research area, but it offers an especially clear example of how compositional complexity can be coupled to electronic and defect engineering. A recent perspective on HEO photocatalysis identified only approximately 38 reported HEO photocatalysis studies at the time of its analysis, indicating that this research field remains substantially less mature than conventional oxide photocatalysis. The same study emphasized the potential of HEOs to tune band structures, generate lattice defects and strain, modify surface chemistry, and improve structural stability. The photocatalytic behavior of HEOs can be understood through a sequential relationship in which HEO composition determines the local structural and defect characteristics, which subsequently influence the electronic structure and photogenerated charge generation and separation, ultimately governing surface reactions and photocatalytic performance.

The importance of composition lies in the fact that different cations can modify band structure and local chemical environments simultaneously. Oxygen vacancies and lattice distortion can additionally introduce localized states or modify carrier transport, while surface cation diversity can alter adsorption and reaction pathways. Therefore, HEO photocatalysis provides an unusually complex but potentially highly tunable materials-design problem [153].

5.2.1. Photocatalytic H_2 Evolution

Edalati et al. reported the first high entropy photocatalyst for hydrogen evolution. Their material had an overall composition of TiHfZrNbTaO11 and consisted of two perovskite-related phases. It possessed a d0 electronic configuration, visible-light absorption with a band gap of approximately 2.9 eV, and band positions considered suitable for water splitting.

The material successfully generated hydrogen photocatalytically, establishing HEOs as a new class of photocatalytic materials (Figure 11a) [164].

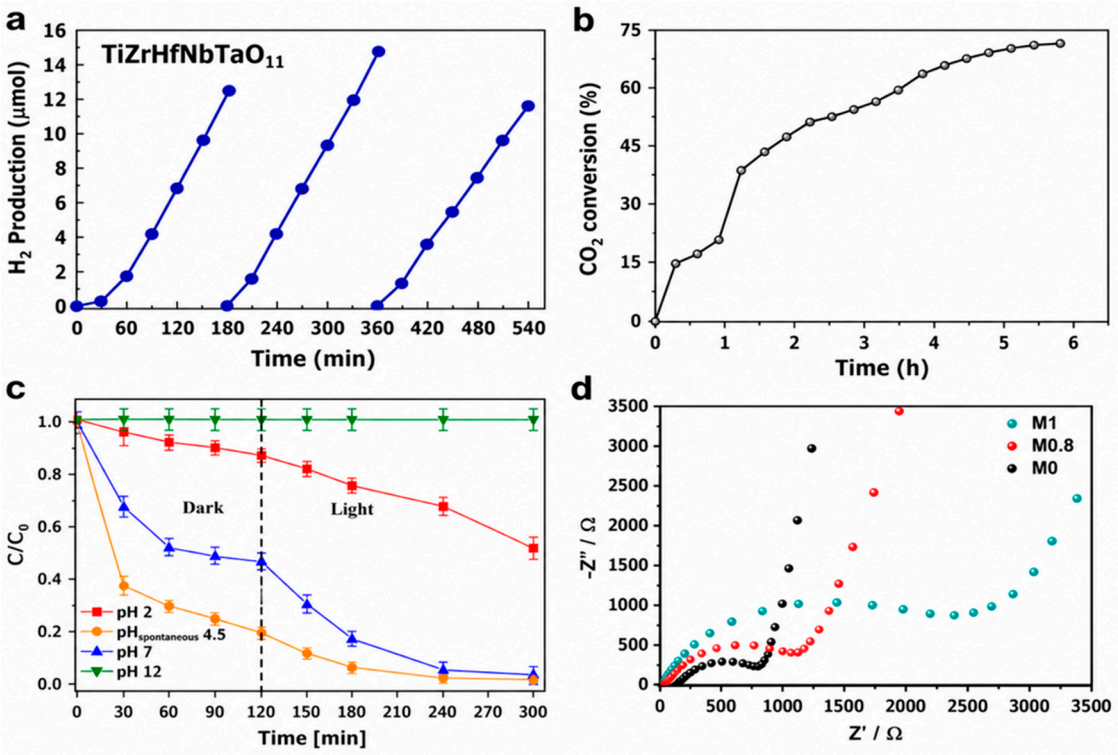


Figure 11. Representative functional characteristics of high-entropy oxides. (**a**) Repeated photocatalytic H_2-production cycles for $TiZrHfNbTaO_{11}$ [164] (Copyright © 2020 Royal Society of Chemistry). (**b**) Time-dependent photocatalytic CO_2 conversion over a fluorite HEO [165] (Copyright © 2025 Elsevier). (**c**) Effect of solution pH on dark adsorption and visible-light photocatalytic degradation of gallic acid, expressed as C/C_0 [118] (Copyright © 2026 Elsevier). (**d**) Nyquist plots for $(Mg_x,Co,Ni,Cu,Zn)O$ anodes designated M1, M0.8, and M0 [166] (Copyright © 2021 Elsevier).

The significance of this study extends beyond demonstrating H_2 production. The selected elements provide complementary functions within a common multicationic oxide framework, illustrating how cation selection can be used to simultaneously influence optical absorption, band alignment, structural stability, and surface redox chemistry. This type of multidimensional optimization is precisely where ML can provide an advantage over isolated experimental screening. At present, however, the HEO photocatalysis literature does not yet support the conclusion that ML can universally predict H_2 evolution rates directly from composition. A more defensible interpretation is that ML can first target proximate descriptors, such as band gap, band-edge characteristics, defect energetics, local electronic structure, or phase stability, which can subsequently be correlated with photocatalytic H_2 production [167].

5.2.2. CO_2 Reduction

HEOs have also demonstrated potential for photocatalytic CO_2 conversion. Chen et al. [165] prepared the fluorite oxide $(Ce_{0.2}Zr_{0.2}Ti_{0.2}Sn_{0.2}Ca_{0.2})O_{2-\delta}$ by solid-state reaction and investigated its photocatalytic CO_2 reduction. Figure 11b shows the time-dependent CO_2 conversion reported in that study. The influence of particle size on photocatalytic performance illustrates why data-driven HEO design should consider processing and microstructure alongside composition. These variables affect surface accessibility and

electronic and defect-related properties, which must be considered when developing models of photocatalytic activity.

Electrochemical CO_2 reduction has also been investigated using entropy-engineered perovskite fuel electrodes in solid oxide electrolysis cells [120]. This is distinct from the photocatalytic process illustrated in Figure 11b. For ML, both application areas motivate condition-specific models of adsorption energies, vacancy formation, electronic structure, and charge transport before direct prediction of product yields.

5.2.3. Pollutant Degradation

The application of HEOs to pollutant removal is expanding rapidly because multi-cationic disorder can produce chemically diverse surface sites and defect-rich structures suitable for advanced oxidation and photocatalytic processes. The recent literature reports HEOs for degradation of organic pollutants under visible light and for persulfate/peroxymonosulfate (PMS)-based advanced oxidation processes.

A particularly informative recent example is the investigation of lattice-distorted HEOs for PMS activation. The study by [73] reported enhanced ROS generation and effective tetracycline removal and used EPR to identify the principal reactive species. Importantly, lattice distortion was linked to both catalytic activity and material stability, rather than being treated as an isolated structural parameter. Mancuso et al. [118] investigated a bixbyite-structured $(CeZrYbErGd)_2O_3+\delta$ HEO photocatalyst for gallic-acid degradation. The material showed visible-light activity, approximately 99% degradation efficiency under the reported optimum conditions, and good recyclability. Figure 11c illustrates the effect of solution pH on dark adsorption and subsequent photocatalytic degradation. The study demonstrates the continued expansion of HEO photocatalysis beyond conventional rock-salt and spinel systems toward additional structural families [73,118].

For ML, pollutant degradation represents a potentially rich but challenging target because the measured removal rate depends not only on the catalyst composition and structure but also on pollutant concentration, pH, oxidant concentration, light intensity, reactor geometry, adsorption equilibrium, and reaction time. Consequently, future ML datasets should include reaction-condition descriptors rather than attempting to predict degradation rates from HEO composition alone.

5.2.4. Nitrogen Fixation and Other Emerging Photocatalytic Reactions

Nitrogen fixation represents an emerging application in which HEO compositional complexity can be used to manipulate both light absorption and charge-carrier behavior. For instance, Nguyen et al. [168] investigated HEO photocatalysts containing d^0 and mixed $d^0 + d^{10}$ cationic configurations for ammonia production from nitrogen and water. The authors found that incorporating d10 cations such as Ga and Zn improved photocatalytic behavior and associated this enhancement with improved light absorption, charge separation, and carrier lifetime. This result provides a useful example of composition-dependent electronic design, in which the electronic configurations of the constituent cations influence light absorption and photogenerated charge-carrier dynamics, thereby affecting the efficiency of photocatalytic N_2 reduction [168].

More broadly, HEOs are being considered for other emerging photocatalytic reactions, such as H_2O_2 production over an octonary HEO [169], because their compositional flexibility can potentially be used to tune redox potentials, adsorption sites, defect populations, and carrier dynamics. However, these emerging applications remain too limited to justify broad claims about universal HEO superiority. The current literature instead supports the view that HEOs constitute an expandable compositional platform for photocatalyst design.

5.3. HEOs for Electrochemical Energy Storage

Beyond catalysis, HEOs have attracted significant interest for electrochemical energy-storage applications because multicationic disorder can provide structural flexibility, multiple redox-active centers, defect tolerance, and compositional control. A recent comprehensive review classified these applications into electrochemical energy-storage devices including lithium-ion and sodium-ion batteries and supercapacitors and electrochemical conversion reactions [18] including HER, OER, ORR and CO_2RR.

For lithium-based systems, spinel and other HEO structures have been investigated as anodes because multiple transition-metal cations can participate in redox processes. For example, (FeCoNiCuZn)3O4 has been investigated as an anode for lithium-ion capacitors, where the HEO exhibited a reversible capacity of approximately 540 mA h g^{-1} after 500 cycles at 1 A g^{-1} [170]. A separate study by Wang et al. [166] investigated (Mgx,Co,Ni,Cu,Zn)O anodes with Mg:Co:Ni:Cu:Zn ratios of 1:1:1:1:1 (M1), 0.8:1:1:1:1 (M0.8), and 0:1:1:1:1 (M0). Figure 11d compares their Nyquist plots and illustrates the dependence of electrochemical impedance on composition. Operando and post-cycling studies further indicate that multication redox and nanoscale metal/oxide reorganization can contribute to the reversible storage mechanism, so nominal composition alone is an incomplete descriptor of the electrochemically active state [85,166,171].

Sodium-ion storage has also become an active area. A 2026 study reported a layered high-entropy oxide cathode, $NaMn_{0.2}Ni_{0.2}Co_{0.2}Fe_{0.2}Al_{0.1}Cu_{0.1}O_2$, for aqueous sodium-ion batteries. The material showed Na^+-dependent charge storage involving predominantly intercalation/deintercalation with a smaller capacitive contribution and delivered 55 mA h g^{-1} in the reported full-cell configuration [172].

More recent work has also demonstrated that configurational entropy can be used to manipulate local coordination environments in sodium-ion cathodes. A 2026 study on entropy-designed layered oxides investigated how configurational entropy affects local crystal-field stabilization and structural robustness, highlighting a direction in which entropy engineering is linked directly to local electronic structure rather than being treated only as a phase-stability descriptor [173].

A particularly recent sodium-ion study further demonstrated a phase-transition-free high-entropy oxide cathode based on fluorine-substituted P2-type material. The reported cathode maintained nearly zero volume variation over the investigated voltage range and achieved 182 mA h g^{-1} at 0.1 C, demonstrating how compositional and local-coordination engineering can be used to address structural degradation during cycling. These examples demonstrate why energy-storage HEOs are promising targets for data-driven design. Their performance depends on multiple coupled variables including redox-active cations, crystal structure, local coordination, phase stability, ionic diffusion, defect chemistry, and cycling stability that can potentially be incorporated into multi-objective ML models [174].

5.4. HEOs for Environmental Applications

HEOs are increasingly being explored for environmental technologies in which chemical stability, redox activity, oxygen mobility, and surface reactivity are important. These applications include pollutant degradation, advanced oxidation, catalytic gas conversion, chemical looping, and potentially environmental sensing. For pollutant degradation and advanced oxidation, lattice distortion and oxygen-vacancy chemistry can promote reactive oxygen species generation while simultaneously providing structural robustness. The recent PMS-activation study demonstrates this relationship directly, showing that lattice distortion can enhance both catalytic activity and durability during tetracycline degradation [73].

CO oxidation represents another environmentally relevant application. The fluorite HEO developed by Riley et al. [148] achieved catalytic activity without relying on platinum-

group metals and was explicitly designed to balance activity, oxygen storage and thermal stability. Chemical looping represents a particularly promising bridge between environmental engineering and data-driven HEO discovery. Brorsson et al. [29] developed an active-learning strategy for high-entropy oxygen carriers in chemical looping by integrating predictive modeling, uncertainty estimation and iterative candidate selection. Their work demonstrated that active learning can focus computational effort on the most promising regions of the composition space while reducing the amount of data required relative to conventional trial-and-error or grid-search approaches; in both studies, candidate performance was evaluated with first-principles-informed calculations and MLIPs, so experimental synthesis of the proposed oxygen carriers remains to be reported. This study is especially important for the present review because it demonstrates a direct data-driven environmental application of high-entropy materials, rather than merely applying ML to a conventional HEO property. For environmental sensing and gas purification, the multicationic composition of HEOs may similarly provide tunable adsorption, redox and transport characteristics. However, the present literature does not yet justify treating these applications as mature areas of ML-guided HEO design. They should therefore be presented as emerging opportunities, particularly for future models incorporating target-gas adsorption energy, surface electronic structure, defect formation, and environmental stability [29,30].

5.5. ML-Guided Prediction of Functional Performance

5.5.1. Cross-Application Performance Benchmarks and Their Use in ML

Application tables are useful only when the reported metric is tied to the experimental protocol. Overpotential, Tafel slope, light-off temperature [24,119,175], H_2 production rate, apparent quantum yield, degradation efficiency [21,118,164], energy demand, current density [119,175], specific capacity, and cycling retention [172,176] quantify different physical processes and cannot be pooled into a single ML target without harmonizing test conditions. The following sections distinguish the validation evidence of representative discovery studies, the relative maturity of application areas, and quantitative model and experimental benchmarks.

For example, the apparent rate constants reported for Pd-catalyzed reduction of 4-nitrophenol and dyes [75] describe an aqueous reduction protocol and should not be pooled with HEO photocatalytic degradation rates. Likewise, methane conversion and product selectivity in plasma reactors require the associated energy input and reactor conditions to be reported [163]. These examples reinforce the need to retain reaction identity, activation mode, and operating conditions as separate dataset fields.

The preceding applications demonstrate that HEO performance depends on a hierarchy of structural and chemical factors. However, the current HEO-ML literature is not equally mature across all levels of this hierarchy. Direct prediction of final application-level metrics such as H_2 evolution rate, CO_2 conversion rate, pollutant degradation rate, catalytic selectivity, or battery lifetime remains considerably less developed than prediction of phase, structure, defect energetics, or intermediate physicochemical descriptors [147].

For example, ML-guided oxygen-vacancy engineering has already been demonstrated for high-entropy perovskite oxides. Tukur et al. [10] developed a graph-neural-network framework for predicting oxygen-vacancy formation energies and used it to screen compositions before experimental electrocatalytic validation. Similarly, ML has been applied to HEO band-gap determination from UV-vis spectroscopy, demonstrating the ability of data-driven methods to accelerate extraction of optical descriptors [108].

These examples suggest that a practical near-term strategy is to adopt a hierarchical prediction framework rather than attempting to directly predict final functional performance from composition alone. In such an approach, ML models sequentially relate composition

to phase formation, local structure, defect chemistry, electronic and reaction descriptors, and ultimately functional performance, thereby preserving the underlying physicochemical relationships that govern HEO behavior.

For photocatalysis, this could involve predicting band gap, band-edge positions, oxygen-vacancy formation energy, work function, adsorption energies, and carrier-related descriptors before experimentally evaluating H_2 or CO_2-reduction activity. For heterogeneous catalysis, the corresponding intermediate variables could include adsorption energies, vacancy formation energies, surface segregation tendencies, oxidation-state distributions, and activation barriers. This hierarchical strategy is likely to be more robust than direct composition-to-performance regression because it preserves the physical causality embedded in the HEO system.

5.5.2. Distinguishing Accurate ML Prediction from ML-Guided Discovery of Improved Catalysts

Accurate prediction of phase stability, crystal structure, or a descriptor does not by itself show that ML has produced a better catalyst. For a critical reading of the literature, four types of evidence can be distinguished; they correspond to combinations of the validation categories defined in Section 6.6. (i) Retrospective prediction, in which ML reproduces known labels or properties within an existing dataset: examples include crystal-structure classification with an accuracy of 0.977 [7], DFT-ML prediction of lattice constants and formation energies of spinel HEOs [143], band-gap extraction from UV-vis spectra [108], and empirical phase maps [97]. (ii) Prospective computational screening, in which ML or MLIPs propose untested compositions that are checked by DFT or against known experiments but are neither synthesized nor tested catalytically: examples include MLIP-based synthesizability screening [8], active-learning searches for oxygen carriers evaluated with first-principles-informed calculations and MLIPs [29,30], and composition optimization based on adsorption-energy distributions [154,155,157]. (iii) Synthesis and structural confirmation of an ML-selected material without a demonstrated functional improvement, as for the fluorite $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$ [94] and the non-equimolar, Ca-containing rock-salt HEO [9]. (iv) Synthesis and catalytic testing of an ML-selected HEO, with performance compared against the training data or against reference catalysts.

To our knowledge, only a few HEO studies currently reach type (iv). Duan et al. [24] used classification and regression models, trained with deliberately collected negative data, to select $Ni_{0.04}Co_{0.48}Zn_{0.36}V_{0.12}Cr_2O_x$ for CH_4 oxidation; the catalyst reached T_{90} = 345 °C and remained stable for more than 7000 h with sulfur and moisture resistance. Nie et al. [11] used a closed active-learning loop to identify four new spinel HEOs, the best of which produced 251 μmol g_{cat}^{-1} min^{-1} H_2 in the water–gas shift reaction at 300 °C, exceeding Pt/γ-Al_2O_3 and Cu/ZnO/Al_2O_3 references under the same conditions, with stable operation for 120 h. Tukur et al. [10] ranked perovskite HEOs by GNN-predicted oxygen-vacancy formation energies and synthesized three compositions whose measured oxygen-vacancy content and OER overpotentials (670, 600, and 540 mV) followed the predicted order. Even in these cases, the improvement is demonstrated within a defined composition family and under specific test conditions; it does not establish that the ML-selected material outperforms the best catalysts reported for the same reaction. No HEO study has yet reached type (iv) for photocatalytic H_2 evolution, photocatalytic CO_2 reduction, or pollutant degradation.

Where ML has added value beyond conventional computational and experimental approaches, it has done so in three identifiable ways: (i) it has made it practical to sample chemical disorder at scale, for example, through MLIP screening of thousands of compositions and structures that would be prohibitively expensive with DFT [8,9]; (ii) it has reduced the number of syntheses needed to find a good catalyst, as in active learning with

only a few experimental rounds [11]; and (iii) it has used negative data and interpretable descriptors to define where catalysts should not be sought [24,94]. Type (iv) evidence should therefore be the reference point when claims of "ML-accelerated catalyst discovery" are evaluated.

5.6. Case Studies of Successful Data-Driven HEO Discovery

The emerging literature can be organized into several levels of data-driven HEO discovery. Table 3 summarizes representative studies and the validation evidence reported for each.

Table 3. Representative recent ML-guided HEO discovery studies and their validation evidence.

Target System	ML/Computational Approach	Key Descriptors/Inputs	Main Result	Validation Evidence	References
Rock-salt HEOs	ML interatomic potential (MLIP)	Bond-length distribution; ΔH_{mix}	Mapped single-phase stability and identified additional candidates	Agreement with all known rock-salt HEOs; new non-equimolar Ca-containing composition synthesized	[9]
HEO crystal structures	XGBoost/supervised ML + SHAP	Anion/cation radius ratio; electronegativity descriptors	High-accuracy multiclass structure prediction; interpretable descriptors	External/held-out model validation	[7]
Fluorite HEOs	Ensemble classification + SHAP	Composition/statistical descriptors; electron affinity; electronegativity	Identified $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$ as a new fluorite HEO	Experimental: XRD + TEM/SAED	[94]
High-entropy perovskites	OxiGraphX graph neural network	Atomic identities; local graph environments	Predicted oxygen-vacancy formation energies for screening	Experimental: three compositions synthesized; OVC and OER trend	[10]
Spinel HEOs	DFT + ML	Composition; structural/site descriptors	Predicted formation energies and lattice constants	Computational validation	[143]
HEO oxygen carriers	Active learning	Predicted performance; uncertainty/acquisition information	Data-efficient exploration of composition space	Computational (first-principles/MLIP-informed evaluation within the AL loop); no synthesis reported	[29,30]
Literature-derived HEOs	LLM data mining + XGBoost/SHAP	Extracted composition, phase and synthesis data; statistical features	Built a structured database and predicted crystal-phase classes	Database/ML validation	[98]
Broad HEO search space	MLIP-based screening	Cation-energy variance; ΔH_{mix}; structural candidates	Expanded screening to larger elemental and structural spaces and addressed synthesizability	Candidate-level computational validation	[8]

Sivak et al. [9] used a machine learning interatomic potential to explore single-phase rock-salt HEOs. The framework combined atomistic ML with physically meaningful descriptors such as bond-length distributions and mixing enthalpy and identified additional candidate compositions beyond the experimentally established materials. This represents one of the clearest examples of ML being used to screen previously unexplored HEO compositions before experimental synthesis.

Dicks et al. [8] subsequently expanded MLIP-based screening to broader elemental and structural spaces and explicitly addressed HEO synthesizability. This represents an

important transition from calculating which compositions are energetically favorable to asking which candidate materials may be realistically accessible experimentally.

Tukur et al. [10] employed DFT-calculated oxygen-vacancy formation energies to train a graph neural network for high-entropy perovskite oxides, which was subsequently used to screen candidate compositions for experimental evaluation. This work demonstrates an integrated composition-defect energetics-candidate screening-experimental validation strategy and provides a representative example of how ML can facilitate defect-informed functional design of HEOs.

A recent study developed an ML-assisted methodology for fluorite HEO design and experimentally synthesized $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$. The model was evaluated using nested stratified cross-validation and external validation, yielding weighted-average F_1-scores of 77% and 70%, respectively. XRD and TEM/SAED characterization supported the predicted fluorite phase, while SHAP analysis identified electron-affinity and electronegativity-related descriptors as important contributors to phase prediction. This study is particularly significant because it demonstrates an integrated workflow in which existing HEO data are used to train an ML classification model, the model identifies a previously unexplored composition, and the prediction is subsequently tested through experimental synthesis and structural characterization, thereby extending ML-assisted HEO research beyond computational screening alone [94].

Brorsson et al. [29] developed an active-learning framework for high-entropy oxygen carriers in which model predictions and uncertainty estimates were used to iteratively select the most informative compositions for further evaluation. By prioritizing candidates based on their expected information gain rather than screening compositions uniformly, the approach enabled more efficient exploration of the high-entropy compositional space than conventional trial-and-error and grid-search strategies. This study illustrates the transition from passive screening, in which predefined candidates are evaluated sequentially, to adaptive discovery, in which each new calculation or experiment is selected based on the information gained from previous results. Table 4 places these results in the context of the relative maturity of data-driven HEO research across functional applications.

Table 4. Relative maturity of data-driven HEO research across major functional application areas.

Application Area	Experimental HEO Performance (Non-ML)	ML Descriptor/Property Prediction	ML-Guided Candidate Selection	Experimental Validation of an ML-Selected HEO	References
Phase stability/synthesizability	Strong	Established	Demonstrated	Demonstrated (structural) [9,94]	[7–9,94,103]
Crystal-structure selection	Strong	Established	Demonstrated	Demonstrated (structural only) [94]	[7,94,97]
OER/HER/ORR	Strong	Emerging	Demonstrated (OER) [10,157]	Demonstrated for OER, one study [10]; not reported for HER/ORR	[10,119,150,157, 158,175]
Thermal catalysis (CO/CH_4 oxidation, water–gas shift)	Strong	Emerging	Demonstrated [11,24]	Demonstrated: CH_4 oxidation [24]; water–gas shift [11]	[11,24,86,148]
Photocatalytic H_2 evolution	Growing	Limited (band-gap extraction only) [108]	Not reported	Not reported	[21,108,164,177]

Table 4. *Cont.*

Application Area	Experimental HEO Performance (Non-ML)	ML Descriptor/Property Prediction	ML-Guided Candidate Selection	Experimental Validation of an ML-Selected HEO	References
CO_2 conversion	Growing	Not reported for HEOs (HEA studies only [154])	Not reported	Not reported	[120,153,165]
Pollutant degradation/AOP	Growing	Not reported	Not reported	Not reported	[73,118]
Energy storage	Strong	Emerging [103,147]	Emerging (computational) [147]	Limited [103]	[18,20,103,147,174]
Oxygen carriers/chemical looping	Growing	Emerging	Demonstrated (computational) [29,30]	Not reported	[29,30]

Classification criteria. Experimental HEO performance (non-ML): Strong = numerous independent experimental studies and dedicated reviews; Growing = an expanding but still limited experimental literature. ML columns: Established = three or more independent HEO-ML studies with reported validation; Demonstrated = at least one peer-reviewed HEO study in which the step was completed for this application; Emerging = ML applied only to proxies, related systems, or computational candidates, without completing the step for this application; Limited = ML used only for data processing or an isolated indirect descriptor; Not reported = no HEO study identified in our literature search. "Experimental validation of an ML-selected HEO" requires that a composition chosen by an ML model was synthesized and its target property measured; the experimental performance of HEOs that were not selected by ML is counted only in the second column. The references column lists the studies supporting the classification.

5.7. Quantitative Benchmarking of Representative HEO-ML Studies

A high headline accuracy is not directly comparable across HEO-ML studies because the scientific effort of the prediction depends on dataset size, class balance [26,131], difference between training and test data [28], target definition, and whether the model is evaluated retrospectively or prospectively. The 0.977 accuracy reported for HEO crystal-structure classification [7], for example, is an impressive internal classification result but addresses a different task from the external-validation F1-score of 0.70 reported before experimental fluorite discovery [94], the ~0.023 eV defect-energy MAE reported for OxiGraphX [10], or the prospective active-learning discovery of HEOs with high H_2 production in the thermocatalytic water–gas shift reaction [11]. Table 5 therefore compares model metrics together with the discovery-space test and experimental validation rather than treating all scores as equivalent measures of progress.

Duan et al. [24] provide a complementary catalysis example in which negative data were deliberately collected, classification and regression models were coupled to screening, and the final candidate was experimentally demonstrated for CH_4 oxidation, with T_{90} = 345 °C, sulfur and moisture resistance, and >7000 h stability. Nie et al. [11] used a prospective active-learning loop to identify four new HEOs and a reported H_2 production rate of 251 μmol g_{cat}^{-1} min^{-1} in the thermocatalytic water–gas shift reaction at 300 °C, which should not be confused with photocatalytic H_2 evolution. Sivak et al. and Dicks et al. instead used MLIPs to move the bottleneck from direct DFT enumeration toward large-scale atomistic screening [8,9]. Collectively, these studies show that the most informative comparison is the complete chain from data construction and representation to uncertainty, candidate selection, chemical-domain testing, and experimental outcome.

Table 5. Quantitative and validation-oriented comparison of representative HEO machine learning studies. Reported metrics are not directly interchangeable because datasets and splitting strategies differ.

HEO System/Target	Model/Strategy	Reported Model Metric (Validation Procedure)	Dataset Size/Discovery-Space Test	Experimental Validation	Key Lesson	Reference
Multiclass HEO crystal structure	XGBoost + SHAP	ACC 0.977; F1 0.975 (3-fold CV; 7:3 train/test split after ADASYN class balancing)	112 experimental compositions (3 classes); held-out test set	No new synthesis in model paper	High internal accuracy; radius/electronegativity descriptors dominate.	[7]
Six-class HEO phase prediction	Ensemble classification + SHAP	Weighted F1 0.77 (nested stratified CV); 0.70 (external validation)	Size NR; external dataset	Yes; $Ce_{0.2}La_{0.2}Nd_{0.2}Mg_{0.2}Al_{0.2}O_{2-\delta}$, XRD + TEM/SAED	External validation exposed a realistic drop from internal CV while retaining discovery utility.	[94]
HEPO oxygen-vacancy formation energy	OxiGraphX GNN + DFT	MAE ≈ 0.023 eV for average OVFE (best of 10 runs; mean 0.025 eV); 10% test split + 3 unseen compounds	9000 DFT OVFEs from 9 compounds; 3 compounds held out	Yes; 3 compositions; OVC and OER order validated	Local graph representations can bridge defect energetics and experiment when the target is mechanistically proximal.	[10]
Entropy-stabilized catalysts for CH_4 oxidation	Classification + regression + screening	Cross-validation score > 0.7 (metric and number of folds NR in the abstract)	Size NR; ACr_2O_x screening + negative data	Yes; CH_4 oxidation: T_{90} = 345 °C, >7000 h, SO_2/H_2O resistant	Negative examples and experimental iteration materially improve discovery utility.	[24]
HEO composition for high H_2 production (water–gas shift)	Active learning closed loop	XGBoost phase classifier + activity regressor; single score NR	514 initial samples; ~14,443 candidates; prospective AL rounds	Yes; four new HEOs; WGS at 300 °C: 251 μmol g_{cat}^{-1} min^{-1} H_2; 120 h	Prospective closed-loop performance is stronger evidence than retrospective accuracy alone.	[11]
Rock-salt HEO stability	MLIP + bond-length/ΔH_{mix} descriptors	Energy MAE 18.7 meV/atom vs. DFT (CHGNet); all 8 known compositions resolved	154 equimolar compositions from 8 cations	Yes; new non-equimolar Ca-containing HEO synthesized	Local structural accommodation complements configurational entropy.	[9]
Tetravalent HEO synthesizability	MLIP + cation-energy variance + ΔH_{mix}	ΔH_{mix} vs. DFT: RMS error 24–47 meV/f.u.; compared with alternative descriptors	3003 compositions × 3 structures; 14 elements	Predicted candidate systems	Atomistic screening scales the search while retaining explicit configurational information.	[8]
Single-phase HEO anode factors	Data-driven statistical/ML analysis	Balanced accuracy ≈0.85 (random forest; CV details in SI)	269 synthesized five-cation compositions	No prospective synthesis in core model	Useful for identifying composition factors, but transferability depends on dataset coverage.	[103]
HEO phase mapping	Empirical biplots	Accuracy 78–100% for selected biplots (no train/test split; small test set)	165 literature compositions	No prospective synthesis	Illustrates how compact datasets can yield optimistic separation; external chemistry remains the harder test.	[97]

NR = not reported in the original study (or not stated in its main text); WGS = water–gas shift. Dataset sizes and validation procedures are taken from the cited studies; OVFE = oxygen-vacancy formation energy; OVC = oxygen-vacancy content; AL = active learning; CV = cross-validation.

6. Challenges, Validation Standards, and Future Perspectives

ML is increasingly useful for HEO phase prediction, compositional screening, defect engineering, and active-learning-guided discovery. Yet reliable data-driven design remains limited by the mismatch between the size of HEO chemical space and the quality and diversity of available data. Future progress should therefore emphasize physically meaningful representations, rigorous validation, uncertainty-aware learning, and systematic computational-experimental feedback rather than algorithmic complexity alone [16].

6.1. Data, Validation, and Extrapolation Challenges

HEO ML is constrained by sparse and chemically biased datasets, inconsistent experimental reporting [32,33], limited negative results [34], and strong dependence of phase formation and performance on synthesis conditions [97]. These issues can produce overfitting when related compositions are divided randomly between training and test sets. Future studies should therefore report chemistry focused validation, external testing where possible, calibrated uncertainty, and an explicit applicability domain [40,41]. Failed synthesis and multiphase outcomes should be retained because they define the decision boundary for synthesizability rather than being discarded as inconvenient observations [34]. Thermodynamic stability must also be distinguished from experimental accessibility, particularly for metastable phases and redox sensitive oxides whose outcome depends on oxygen chemical potential and thermal response [35,141].

6.2. Representing HEO Disorder and Physics-Informed ML

Nominal composition does not uniquely define an HEO because chemically equivalent crystallographic sites can possess different local environments, short-range correlations [22,57,59], defect configurations, oxidation states, and sublattice occupancies [61,66]. Structure oriented models such as graph neural networks and MLIPs therefore offer an important route toward representing atomic interactions more explicitly [39,109]. Their transferability across structural families, oxygen stoichiometries, magnetic states, and reconstructed surfaces remains an open problem [38,102]. Physics-informed ML can complement these representations by imposing charge neutrality, composition normalization, thermodynamic consistency, crystallographic compatibility, site-occupancy rules, and defect chemistry, thereby reducing the effective hypothesis space when datasets are small [25,38,102].

6.3. From Predictive ML to Generative and Active Design

The next step beyond predicting known HEOs is to select or generate compositions that are most informative or most likely to satisfy a predefined multi-objective target. Active learning has already demonstrated direct value in HEO discovery [11,29,30], while broader materials studies show how Bayesian optimization [44,45,178] and autonomous experimentation [46–48] can reduce the number of costly evaluations required to identify promising candidates. Generative models could extend this framework by proposing compositions subject to explicit constraints on phase stability, oxidation state balance, site compatibility, defect chemistry, elemental distribution, cost, and synthesizability [179,180]. Fully generative HEO design remains immature and should be distinguished from model-assisted candidate selection. Generated candidates still require physical screening and experimental verification [6,25,34,179].

6.4. Toward Autonomous and Multimodal HEO Discovery

A longer-term objective is to integrate active learning with automated synthesis, characterization, and multimodal scientific data. Such a platform could combine literature, composition [32,33,99], crystal structure, DFT [137,181], spectroscopy, microscopy [137,138], processing conditions, and functional measurements within a continuously updated database [47,182]. Operando HEO studies already demonstrate why working-state information matters. Lithiation and catalytic redox mechanisms can differ substantially from the ex situ structure, and site-specific spectroscopy can identify which cations participate in the reaction [85,86,162,166]. Digital twins and foundation models may eventually support this integration, but sufficiently standardized multimodal HEO datasets do not yet exist. The near-term priority should therefore be domain-specific, uncertainty sensitive systems that

connect model predictions with experimentally reproducible synthesis and characterization rather than autonomous complexity for its own sake [25,102,138,183].

6.5. Research Roadmap for Reliable and Autonomous HEO Discovery

A useful roadmap separates near-term needs for reliable data and validation from medium-term disorder-aware physics and longer-term closed-loop discovery. In the near term, the highest return is likely to come from curated datasets, negative results [32–34], chemistry-oriented splitting, calibrated uncertainty [28,40,41], and preservation of synthesis data [43,142]. Medium-term progress should focus on configurational ensembles, SQS/Monte Carlo sampling [57,83], MLIPs [8,9,93], multimodal characterization [82,85,166], and condition-related performance models [73]. Longer-term autonomous workflows should be judged by experimentally validated candidates, reduced experimental burden, quantified resource savings, and reproducibility rather than by the mere presence of robotics or generative AI [11,29,30,183]. Table 6 organizes these priorities by time horizon, recommended methodological advances, and criteria for success.

Table 6. Research roadmap for reliable, disorder-aware, and progressively autonomous HEO discovery.

Time Horizon	Priority	Recommended Methodological Advance	Success Criterion	References
Near term (1–3 y)	Reliable data and validation	Curated HEO datasets; failed experiments; synthesis metadata; chemistry-aware splits; calibrated uncertainty	Models retain performance on external chemistry and report applicability domains	[28,32,33]
Medium term (3–5 y)	Disorder-aware physics	Configurational ensembles; SQS/Monte Carlo + MLIP; multimodal XAS/PDF/STEM descriptors; process-aware synthesizability	Predictions distinguish same-composition samples with different local states	[9,57,83,84,89]
Medium term (3–5 y)	Condition-aware performance	Couple intrinsic descriptors with reaction/cell variables; harmonized benchmarks	Transfer across laboratories and operating windows	[118–120,183]
Long term (>5 y)	Closed-loop discovery	Active learning + robotics + in situ/operando characterization + multi-objective optimization	Experimentally validated candidates found with fewer experiments and quantified resource savings	[11,30,31,140,183]
Long term (>5 y)	Autonomous knowledge integration	Foundation models/LLM extraction linked to structured databases and physics constraints	New hypotheses are traceable, chemically feasible, and experimentally reproducible	[98,136,138,181,184]

6.6. Minimum Reporting Standard for ML-Ready HEO Research

The field would benefit from a minimum reporting standard that allows experimental observations to be reused without reconstructing critical data from plots or supplementary files. This requirement follows directly from the data curation [32–34] and extrapolation problems [25,28,97] identified in HEO-specific studies and broader materials informatics. A reusable HEO record should separate nominal from measured composition, retain oxygen non-stoichiometry and compositional uncertainty where available, and record synthesis and thermal properties as structured variables rather than free text annotations. Structural labels should distinguish average phase identity from local information such as site occupancy [61,66,82], oxidation state, short-range order, and vacancy indicators [22,57,59].

For model evaluation, random cross-validation should be accompanied by at least one chemistry related variable whenever the stated goal is discovery [25,28] and domain should be flagged explicitly [40,41]. The dataset, feature definitions, splitting code, hyperparameters, seeds, and candidate-selection logic ought to be disclosed [32,33,96,123] where possible so that apparent improvements can be separated from changes in data evaluation protocol [26,28,131].

A practical HEO validation framework can therefore be organized into complementary categories of evidence rather than a single sequence: (A) statistical validation, comprising internal repeated or nested cross-validation and chemical-domain or temporal holdout [25,28]; (B) computational verification of the selected candidates by DFT or a validated MLIP [8,9]; (C) experimental realization, comprising independent synthesis together with structural and local-chemical confirmation [59,94]; (D) functional validation, comprising standardized performance and stability testing under fully reported conditions; and (E) reproducibility, through independent batch or laboratory replication [31,34]. These categories are neither strictly cumulative nor sequential. An active-learning study may proceed directly from model ranking to synthesis and catalytic testing without DFT verification [11,24], whereas an MLIP study may provide extensive DFT verification without any synthesis [8]. Each category answers a different question (whether the model generalizes, whether the predicted energetics are physically sound, whether the material can be made, whether it performs, and whether the result is reproducible), so one category cannot substitute for another. Claims of inverse, autonomous, or closed-loop discovery should state explicitly which categories were satisfied [34], and claims of improved catalysts should be supported at least by categories C and D. This framework intentionally values prospective experimental confirmation more strongly than marginal improvements in retrospective model accuracy. Table 7 summarizes the minimum reporting fields needed to support these validation categories and the reuse of HEO data.

Table 7. Proposed minimum reporting standard for reusable, ML-ready HEO data.

Data Block	Minimum Recommended Fields	Why It Matters for HEO ML	References
Composition	Nominal and measured cation fractions; oxygen stoichiometry; uncertainty	Separates intended formulation from actual chemistry and charge compensation	[35,53,103]
Structure	Phase(s), phase fraction, lattice parameters, site occupancy where resolved, crystallite size	Prevents “single phase” from being treated as a complete structural descriptor	[22,59,61,66]
Local chemistry	Oxidation states, vacancy indicators, local coordination, SRO evidence	Captures chemically distinct states at the same nominal composition	[22,57,59,84]
Synthesis	Precursors, route, pH/solvent, temperature, time, atmosphere, heating/cooling rate, pressure	Synthesizability and defect populations are process-dependent	[8,33,35,131]
Morphology/texture	Particle size, BET area, porosity, support/loading where relevant	Separates intrinsic chemistry from surface-area and dispersion effects	[14,71,74]
Measurement conditions	Reaction/cell conditions, normalization basis, equipment-relevant metadata	Required for cross-laboratory performance comparison	[66,85,86]
Target and uncertainty	Mean value, error/replicate information, detection limits	Enables uncertainty-aware models and avoids false precision	[33,40,41,119,120]

Table 7. *Cont.*

Data Block	Minimum Recommended Fields	Why It Matters for HEO ML	References
Negative results	Failed phase formation, poor performance, decomposition, unsuccessful synthesis	Constrains decision boundaries and reduces publication bias	[34]
ML provenance	Dataset version, split strategy, feature definitions, hyperparameters, code/model version	Required for reproducibility and auditability	[32,33,96]
Applicability domain	Composition/structure distance, uncertainty or conformal interval	Distinguishes interpolation from unsupported extrapolation	[25,28,40,41]
Validation	External chemistry, DFT/MLIP verification, synthesis, structure, function, replication	Makes "discovery" claims comparable across studies	[11,24,25,94]

7. Conclusions

HEOs provide an unusually demanding but scientifically valuable test case for materials informatics because nominal composition, local chemical disorder, defects, electronic structure, and processing history are strongly coupled. The literature now demonstrates credible ML capability for phase and structure classification, MLIP-enabled composition screening, and oxygen-vacancy prediction. Assessed with the complementary validation categories proposed here, most studies provide statistical or computational evidence (categories A and B), and experimental realization of ML-selected HEOs (category C) has been reported in several cases. Functional validation of ML-selected catalysts (category D) has so far been achieved for only a few reactions, notably CH_4 oxidation, the water–gas shift reaction, and the OER, and not yet for photocatalytic H_2 evolution or CO_2 conversion, while independent replication (category E) remains rare. Lessons from high-entropy alloy catalysis indicate that the most informative ML targets for catalysis are site-resolved distributions of adsorption and vacancy-formation energies that can be coupled to kinetic models, rather than composition-averaged descriptors. The next challenge is not simply to increase algorithmic complexity but to improve the physical fidelity of representations and the rigor with which extrapolation is tested. Composition-only models remain useful for rapid screening, but discovery-oriented models increasingly require local-structure, defect, synthesis, and uncertainty information. Direct composition-to-performance prediction is still less mature than phase, structure, or defect prediction because application metrics are strongly conditioned by experimental protocol. High-impact progress will therefore require standardized ML-ready HEO records, chemistry-aware validation, calibrated uncertainty, negative data, and systematic comparison between model-selected candidates and independently synthesized materials. Integration of DFT, disorder-aware MLIPs, graph models, multimodal characterization, active learning, and automated experimentation can ultimately convert HEO discovery from sparse empirical sampling into an iterative, self-improving discovery process. The most credible route toward inverse and autonomous HEO design is consequently a physically constrained and experimentally accountable one in which each stage—representation, prediction, synthesis, structural verification, and functional testing—has an explicit validation criterion.

Author Contributions: Conceptualization, M.B. and M.S.; methodology, M.B.; validation, M.B. and M.S.; formal analysis, M.B., M.S., N.W., Y.Y. and T.R.; investigation, M.B., M.S. and H.S.; resources, N.W.; data curation, M.S., M.B. and H.S.; writing—original draft preparation, M.B. and M.S.; writing—review and editing, N.W., Y.Y. and T.R.; visualization, M.B. and M.S.; supervision, N.W.; project administration, N.W. and Y.Y.; funding acquisition, T.R. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Qassim University (QU-APC-2026).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University (https://www.qu.edu.sa) for financial support (QU-APC-2026).

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Rost, C.M.; Sachet, E.; Borman, T.; Moballegh, A.; Dickey, E.C.; Hou, D.; Jones, J.L.; Curtarolo, S.; Maria, J.-P. Entropy-stabilized oxides. *Nat. Commun.* **2015**, *6*, 8485. [CrossRef] [PubMed]
2. Zou, J.; Tang, L.; He, W.; Zhang, X. High-Entropy Oxides: Pioneering the Future of Multifunctional Materials. *ACS Nano* **2024**, *18*, 34492–34530. [CrossRef] [PubMed]
3. He, Q.; Li, J.; Liu, W.; Zhang, J.; Wang, K.; Liu, J.; Hui, J.; Zhang, H. High entropy oxides for electrochemical energy storage and conversion: A critical review. *J. Power Sources* **2024**, *619*, 235207. [CrossRef]
4. Aamlid, S.S.; Oudah, M.; Rottler, J.; Hallas, A.M. Understanding the Role of Entropy in High Entropy Oxides. *J. Am. Chem. Soc.* **2023**, *145*, 5991–6006. [CrossRef] [PubMed]
5. Kotsonis, G.N.; Almishal, S.S.I.; Marques dos Santos Vieira, F.; Crespi, V.H.; Dabo, I.; Rost, C.M.; Maria, J.P. High-entropy oxides: Harnessing crystalline disorder for emergent functionality. *J. Am. Ceram. Soc.* **2023**, *106*, 5587–5611. [CrossRef]
6. Xiang, H.; Xing, Y.; Dai, F.-Z.; Wang, H.; Su, L.; Miao, L.; Zhang, G.; Wang, Y.; Qi, X.; Yao, L.; et al. High-entropy ceramics: Present status, challenges, and a look forward. *J. Adv. Ceram.* **2021**, *10*, 385–441. [CrossRef]
7. Liu, J.; Wang, A.; Gao, P.; Bai, R.; Liu, J.; Du, B.; Fang, C. Machine learning-based crystal structure prediction for high-entropy oxide ceramics. *J. Am. Ceram. Soc.* **2024**, *107*, 1361–1371. [CrossRef]
8. Dicks, O.A.; Aamlid, S.S.; Hallas, A.M.; Rottler, J. Expanding the search space of high entropy oxides and predicting synthesizability using machine learning interatomic potentials. *Comput. Mater. Sci.* **2026**, *267*, 114581. [CrossRef]
9. Sivak, J.T.; Almishal, S.S.I.; Caucci, M.K.; Tan, Y.; Srikanth, D.; Petruska, J.; Furst, M.; Chen, L.-Q.; Rost, C.M.; Maria, J.-P.; et al. Discovering High-Entropy Oxides with a Machine-Learning Interatomic Potential. *Phys. Rev. Lett.* **2025**, *134*, 216101. [CrossRef] [PubMed]
10. Tukur, P.; Wei, Y.; Zhang, Y.; Chen, H.; Lin, Y.; He, S.; Mo, Y.; Wei, J. Machine Learning-Guided Discovery of High-Entropy Perovskite Oxide Electrocatalysts via Oxygen Vacancy Engineering. *Small* **2025**, *21*, 2501946. [CrossRef] [PubMed]
11. Nie, S.; Xiang, Y.; Wu, L.; Lin, G.; Liu, Q.; Chu, S.; Wang, X. Active Learning Guided Discovery of High Entropy Oxides Featuring High H_2-production. *J. Am. Chem. Soc.* **2024**, *146*, 29325–29334. [CrossRef] [PubMed]
12. Kim, Y.; Kim, H.S.; Kim, H.; Ha, J.; Kang, D.; Lee, J.W. Recent Strategies of Oxygen Carrier Design in Chemical Looping Processes for Inherent CO_2 Capture and Utilization. *Korean J. Chem. Eng.* **2025**, *42*, 3081–3100. [CrossRef]
13. Wang, J.; Zhang, Y. Artificial intelligence in high-entropy materials. *Next Mater.* **2025**, *9*, 100993. [CrossRef]
14. Wu, M.; Liu, L.; Wang, Y.; Zhang, S.; Liu, S.; Low, J.; Fu, F.; Wan, J. Recent advances in high-entropy oxides: Composition modulation, morphology control, and catalytic applications. *Green Energy Environ.* **2026**, *in press*. [CrossRef]
15. He, Y.; Cheng, H.; Zhou, T.; Sun, T.; Liu, Q.; He, T. Artificial Intelligence-Assisted Design of High-Entropy Oxide Catalysts. *Adv. Mater.* **2026**, *38*, e74305. [CrossRef] [PubMed]
16. Mashhadimoslem, H.; Karimi, P.; Elkamel, A.; Yu, A. Toward high entropy material discovery for energy applications using computational and machine learning methods. *npj Comput. Mater.* **2026**, *12*, 50. [CrossRef]
17. Akrami, S.; Edalati, P.; Fuji, M.; Edalati, K. High-entropy ceramics: Review of principles, production and applications. *Mater. Sci. Eng. R Rep.* **2021**, *146*, 100644. [CrossRef]
18. Dong, S.; Ren, R.; Zhang, J.; Bao, X.; Liu, X.; Shi, Q.; Chen, Z.; Shao, H. High-entropy oxides: Emergent materials for electrochemical energy storage and conversion. *J. Mater. Sci. Technol.* **2025**, *227*, 192–215. [CrossRef]
19. Fu, Z.; Huang, P.; Wang, X.; Liu, W.D.; Kong, L.; Chen, K.; Li, J.; Chen, Y. Artificial Intelligence-Assisted Ultrafast High-Throughput Screening of High-Entropy Hydrogen Evolution Reaction Catalysts. *Adv. Energy Mater.* **2025**, *15*, 2500744. [CrossRef]
20. Ran, B.; Li, H.; Cheng, R.; Yang, Z.; Zhong, Y.; Qin, Y.; Yang, C.; Fu, C. High-Entropy Oxides for Rechargeable Batteries. *Adv. Sci.* **2024**, *11*, 2401034. [CrossRef] [PubMed]
21. Tsubota, H.; Jitianu, A.; Kawamura, G. Recent Advances in High-Entropy Oxides for Photocatalytic Applications. *ACS Mater. Lett.* **2025**, *7*, 1042–1056. [CrossRef]
22. Rost, C.M.; Rak, Z.; Brenner, D.W.; Maria, J.P. Local structure of the MgxNixCoxCuxZnxO (x = 0.2) entropy-stabilized oxide: An EXAFS study. *J. Am. Ceram. Soc.* **2017**, *100*, 2732–2738. [CrossRef]

23. Dupuy, A.D.; Chellali, M.R.; Hahn, H.; Schoenung, J.M. Multiscale phase homogeneity in bulk nanocrystalline high entropy oxides. *J. Eur. Ceram. Soc.* **2021**, *41*, 4850–4858. [CrossRef]
24. Duan, X.; Li, Y.; Zhao, J.; Zhang, M.; Wang, X.; Zhang, L.; Ma, X.; Qu, Y.; Zhang, P. Machine Learning Accelerated Discovery of Entropy-Stabilized Oxide Catalysts for Catalytic Oxidation. *J. Am. Chem. Soc.* **2025**, *147*, 651–661. [CrossRef] [PubMed]
25. Riebesell, J.; Goodall, R.E.A.; Benner, P.; Chiang, Y.; Deng, B.; Ceder, G.; Asta, M.; Lee, A.A.; Jain, A.; Persson, K.A. A framework to evaluate machine learning crystal stability predictions. *Nat. Mach. Intell.* **2025**, *7*, 836–847. Correction in *Nat. Mach. Intell.* **2025**, *7*, 1586. [CrossRef]
26. Saal, J.E.; Oliynyk, A.O.; Meredig, B. Machine Learning in Materials Discovery: Confirmed Predictions and Their Underlying Approaches. *Annu. Rev. Mater. Res.* **2020**, *50*, 49–69. [CrossRef]
27. Kaufmann, K.; Maryanovsky, D.; Mellor, W.M.; Zhu, C.; Rosengarten, A.S.; Harrington, T.J.; Oses, C.; Toher, C.; Curtarolo, S.; Vecchio, K.S. Discovery of high-entropy ceramics via machine learning. *npj Comput. Mater.* **2020**, *6*, 42. [CrossRef]
28. Li, K.; Rubungo, A.N.; Lei, X.; Persaud, D.; Choudhary, K.; DeCost, B.; Dieng, A.B.; Hattrick-Simpers, J. Probing out-of-distribution generalization in machine learning for materials. *Commun. Mater.* **2025**, *6*, 9. [CrossRef]
29. Brorsson, J.; Klein Moberg, H.; Hildingsson, J.; Gastaldi, J.; Mattisson, T.; Hellman, A. Data-Efficient Design of High-Entropy Oxygen Carriers for Chemical Looping Using Active Learning. *ACS Mater. Au* **2026**, *6*, 319–326. [CrossRef] [PubMed]
30. Brorsson, J.; Klein Moberg, H.; Gastaldi, J.; Hildingsson, J.; Mattisson, T.; Hellman, A. Discovery of high-entropy perovskite oxygen carriers for chemical looping applications via an autonomous active learning protocol. *Mater. Today Energy* **2026**, *57*, 102239. [CrossRef]
31. Tom, G.; Schmid, S.P.; Baird, S.G.; Cao, Y.; Darvish, K.; Hao, H.; Lo, S.; Pablo-García, S.; Rajaonson, E.M.; Skreta, M.; et al. Self-Driving Laboratories for Chemistry and Materials Science. *Chem. Rev.* **2024**, *124*, 9633–9732. [CrossRef] [PubMed]
32. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef] [PubMed]
33. Hart, M.; Idanwekhai, K.; Alves, V.M.; Miller, A.J.M.; Dempsey, J.L.; Cahoon, J.F.; Chen, C.-H.; Winkler, D.A.; Muratov, E.N.; Tropsha, A. Trust Not Verify? The Critical Need for Data Curation Standards in Materials Informatics. *Chem. Mater.* **2024**, *36*, 9046–9055. [CrossRef]
34. Raccuglia, P.; Elbert, K.C.; Adler, P.D.F.; Falk, C.; Wenny, M.B.; Mollo, A.; Zeller, M.; Friedler, S.A.; Schrier, J.; Norquist, A.J. Machine-learning-assisted materials discovery using failed experiments. *Nature* **2016**, *533*, 73–76. [CrossRef] [PubMed]
35. Almishal, S.S.I.; Furst, M.; Tan, Y.; Sivak, J.T.; Bejger, G.; Petruska, J.; Ayyagari, S.V.G.; Srikanth, D.; Alem, N.; Rost, C.M.; et al. Thermodynamics-inspired high-entropy oxide synthesis. *Nat. Commun.* **2025**, *16*, 8211. [CrossRef] [PubMed]
36. Sutariya, B.; Sarkar, P.; Indurkar, P.D.; Karan, S. Machine learning-assisted performance prediction from the synthesis conditions of nanofiltration membranes. *Sep. Purif. Technol.* **2025**, *354*, 128960. [CrossRef]
37. Schütt, K.T.; Kindermans, P.-J.; Sauceda Felix, H.E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. In *Advances in Neural Information Processing Systems*; Guyon, I., von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates: Red Hook, NY, USA, 2017; Volume 30.
38. Deng, B.; Zhong, P.; Jun, K.; Riebesell, J.; Han, K.; Bartel, C.J.; Ceder, G. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat. Mach. Intell.* **2023**, *5*, 1031–1041. [CrossRef]
39. Chen, C.; Ong, S.P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2022**, *2*, 718–728. [CrossRef] [PubMed]
40. Gawlikowski, J.; Tassi, C.R.N.; Ali, M.; Lee, J.; Humt, M.; Feng, J.; Kruspe, A.; Triebel, R.; Jung, P.; Roscher, R.; et al. A survey of uncertainty in deep neural networks. *Artif. Intell. Rev.* **2023**, *56*, 1513–1589. [CrossRef]
41. Angelopoulos, A.N.; Bates, S. Conformal Prediction: A Gentle Introduction. *Found. Trends Mach. Learn.* **2023**, *16*, 494–591. [CrossRef]
42. Tshitoyan, V.; Dagdelen, J.; Weston, L.; Dunn, A.; Rong, Z.; Kononova, O.; Persson, K.A.; Ceder, G.; Jain, A. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **2019**, *571*, 95–98. [CrossRef] [PubMed]
43. Kononova, O.; Huo, H.; He, T.; Rong, Z.; Botari, T.; Sun, W.; Tshitoyan, V.; Ceder, G. Text-mined dataset of inorganic materials synthesis recipes. *Sci. Data* **2019**, *6*, 203. Correction in *Sci. Data* **2019**, *6*, 273. [CrossRef] [PubMed]
44. Häse, F.; Roch, L.M.; Kreisbeck, C.; Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **2018**, *4*, 1134–1145. [CrossRef] [PubMed]
45. Lookman, T.; Balachandran, P.V.; Xue, D.; Yuan, R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput. Mater.* **2019**, *5*, 21. [CrossRef]
46. Stach, E.; DeCost, B.; Kusne, A.G.; Hattrick-Simpers, J.; Brown, K.A.; Reyes, K.G.; Schrier, J.; Billinge, S.; Buonassisi, T.; Foster, I.; et al. Autonomous experimentation systems for materials development: A community perspective. *Matter* **2021**, *4*, 2702–2726. [CrossRef]

47. Häse, F.; Roch, L.M.; Aspuru-Guzik, A. Next-Generation Experimentation with Self-Driving Laboratories. *Trends Chem*. **2019**, *1*, 282–291. [CrossRef]
48. Burger, B.; Maffettone, P.M.; Gusev, V.V.; Aitchison, C.M.; Bai, Y.; Wang, X.; Li, X.; Alston, B.M.; Li, B.; Clowes, R.; et al. A mobile robotic chemist. *Nature* **2020**, *583*, 237–241. [CrossRef] [PubMed]
49. Gong, P.; Shen, Y.; Sun, X.; Ma, M.; Zhang, R.; He, F.; Fang, D.; Xie, J. High-entropy oxides for catalytic applications: From synthesis to active-site modulation and computational insights. *Fuel* **2026**, *410*, 137931. [CrossRef]
50. Liu, C.; Li, S.; Zheng, Y.; Xu, M.; Su, H.; Miao, X.; Liu, Y.; Zhou, Z.; Qi, J.; Yang, B.; et al. Advances in high entropy oxides: Synthesis, structure, properties and beyond. *Prog. Mater. Sci*. **2025**, *148*, 101385. [CrossRef]
51. Yeh, J.W.; Chen, S.K.; Lin, S.J.; Gan, J.Y.; Chin, T.S.; Shun, T.T.; Tsau, C.H.; Chang, S.Y. Nanostructured High-Entropy Alloys with Multiple Principal Elements: Novel Alloy Design Concepts and Outcomes. *Adv. Eng. Mater.* **2004**, *6*, 299–303. [CrossRef]
52. Cantor, B.; Chang, I.T.H.; Knight, P.; Vincent, A.J.B. Microstructural development in equiatomic multicomponent alloys. *Mater. Sci. Eng. A* **2004**, *375–377*, 213–218. [CrossRef]
53. Manchón-Gordón, A.F.; Panadero-Medianero, P.; Blázquez, J.S. Descriptors for Predicting Single- and Multi-Phase Formation in High-Entropy Oxides: A Unified Framework Approach. *Materials* **2025**, *18*, 3862. [CrossRef] [PubMed]
54. Sen, S.; Palabathuni, M.; Ryan, K.M.; Singh, S. High Entropy Oxides: Mapping the Landscape from Fundamentals to Future Vistas. *ACS Energy Lett*. **2024**, *9*, 3694–3718. [CrossRef] [PubMed]
55. Tiwtusthada, S.; Saelee, T.; Khajondetchairit, P.; Rittiruam, M.; Praserthdam, P.; Praserthdam, S. The development of configurational entropy formulation for high-entropy oxides: Review and outlook. *J. Alloys Compd*. **2025**, *1037*, 182599. [CrossRef]
56. Fracchia, M.; Coduri, M.; Ghigna, P.; Anselmi-Tamburini, U. Phase stability of high entropy oxides: A critical review. *J. Eur. Ceram. Soc*. **2024**, *44*, 585–594. [CrossRef]
57. Aamlid, S.S.; Mugiraneza, S.; González-Rivas, M.U.; King, G.; Hallas, A.M.; Rottler, J. Short-Range Order and Local Distortions in Entropy Stabilized Oxides. *Chem. Mater.* **2024**, *36*, 9636–9645. [CrossRef]
58. Sheikh, K.A.; Khan, M.M. Structural insights and electrochemical applications of high-entropy oxides: A comprehensive review. *Prog. Eng. Sci*. **2025**, *2*, 100105. [CrossRef]
59. Barber, J.P.; Deary, W.J.; Titus, A.N.; Bejger, G.R.; Almishal, S.S.I.; Rost, C.M. Disorder by design: Unveiling local structure and functional insights in high entropy oxides. *Mater. Horiz*. **2025**, *12*, 10453–10477. [CrossRef] [PubMed]
60. Liu, R.-Q.; Wang, Z.-G.; Xie, B.-W.; Hu, L.-Z.; Yu, J.-L.; Cheng, L.; Chang, Z.-D.; Zhang, X.-C.; Wang, H.-C.; Zhang, Y.-H.; et al. Novel high-entropy fluorite oxides with superior phase stability up to 1723 K: Promising candidates for advanced thermal barrier coatings. *Ceram. Int*. **2026**, *52*, 20397–20407. [CrossRef]
61. Bilovol, V.; Szum, M.; Sikora, M.; Berent, K.; Nowakowska, M.; Pitala, K.; Dąbrowa, J.; Cieślak, J. Understanding Cation Distribution in Nominal High-Entropy Oxide Spinels: An Experimental Study. *Chem. Mater.* **2026**, *38*, 5889–5902. [CrossRef]
62. Xiao, M.; Liu, Z.; Di, H.; Bai, Y.; Yang, G.; Medvedev, D.A.; Luo, Z.; Wang, W.; Zhou, W.; Ran, R.; et al. High-entropy materials for solid oxide cells: Synthesis, applications, and prospects. *J. Energy Chem*. **2025**, *104*, 268–296. [CrossRef]
63. Ma, J.; Liu, T.; Ye, W.; He, Q.; Chen, K. High-entropy perovskite oxides for energy materials: A review. *J. Energy Storage* **2024**, *90*, 111890. [CrossRef]
64. Cao, J.; Wu, S.; He, J.; Zhou, Y.; Ma, P. Research progress of high-entropy perovskite oxides in energy and environmental applications: A review. *Particuology* **2024**, *95*, 62–81. [CrossRef]
65. Vayer, F.; Decorse, C.; Bérardan, D.; Dragoe, D.; Dragoe, N. Investigation of the chemical versatility in high-entropy pyrochlores. *J. Am. Ceram. Soc*. **2023**, *106*, 2601–2621. [CrossRef]
66. Masina, S.M.; Nkala, G.C.; Kiefer, M.A.; Cao, C.L.; Nagle-Cocco, L.A.V.; Mhlana, Y.; Stone, K.H. A Multi-Pronged Approach to the Determination of Cation Distribution, Site Selectivity, and Chemical Ordering in "High Entropy" Spinel Oxides. *J. Am. Chem. Soc*. **2025**, *147*, 38283–38290. [CrossRef] [PubMed]
67. Yang, Z.; Xiang, X.; Yang, J.; Zhao, Z.-Y. High-entropy oxides as energy materials: From complexity to rational design. *Mater. Futur.* **2024**, *3*, 042103. [CrossRef]
68. Potter, A.; Wang, Y.; Kong, D.; Li, Y.; Qin, J.; Zheng, X. A Statistical Understanding of Oxygen Vacancies in Distorted High-Entropy Oxides. *ECS Meet. Abstr.* **2025**, *MA2025–02*, 2286. [CrossRef]
69. Verhage, M.; van den Broek, S.J.H.M.; Weijtens, C.; Flipse, C.F.J. Oxygen Vacancies Can Drive Surface Transformation of High-Entropy Perovskite Oxide for the Oxygen Evolution Reaction as Probed with Scanning Probe Microscopy. *ACS Appl. Mater. Interfaces* **2025**, *17*, 23237–23248. [CrossRef] [PubMed]
70. Siniard, K.; Fan, J.; Li, M.; Wang, Q.; Ivanov, A.S.; Wang, T.; Dai, S. A General Strategy for Bandgap Engineering Via Anion-Lattice Doping in High-Entropy Oxides. *Adv. Sci*. **2025**, *12*, e05789. [CrossRef] [PubMed]
71. Lyu, Z.; Wang, Y.; Sun, Y.; Dai, Y. Nanostructured high-entropy oxides for catalysis: Linking entropy to function. *Nanoscale* **2026**, *18*, 7825–7848. [CrossRef] [PubMed]
72. Yang, J.; Gong, S.; Liu, S.; Wu, X.; Guo, Z.; Zhao, Y.; Chen, C. A comprehensive review of high entropy oxides electrocatalysts: Advances, challenges, and prospects. *Nano Res*. **2025**, *18*, 94907774. [CrossRef]

73. Guan, Y.; Zhou, G.; Jiang, Y.; Dong, J.; Li, L.; Yin, J.; Huang, S.; Zhang, L.; Ang, E.H. Lattice distortion effects in high-entropy oxides: Boosting PMS activation for effective and durable pollutant degradation. *Sep. Purif. Technol.* **2025**, *358*, 130267. [CrossRef]
74. Algan Simsek, I.B.; Badr, H.O.; Cardoza, N.; Colin-Ulloa, E.; Schwenk, G.R.; Sudhakar, K.; Wiedwald, U.; Farle, M.; Kalra, V.; Ibrahim, M.A.; et al. Spinel-Structured High Entropy Oxides: Low Temperature Synthesis, Characterization, and Potential Applications. *ACS Omega* **2025**, *10*, 39530–39539. [CrossRef] [PubMed]
75. Ismail, S.; Ahmed, W.; Qadri, A.N.; Wali, N.; Ur Rehman, N.; Yildirim, Y.; Radwan, T. Atmospheric-pressure microplasma synthesis of palladium nanoparticles for environmental catalysis. *Catalysts* **2026**, *16*, 913. [CrossRef]
76. Wali, N.; Ur Rehman, N.; Qadri, A.N.; Gao, F.; Yildirim, Y.; Radwan, T. Plasma-assisted trimetallic nanoparticles and nanoalloys: A review of synthesis routes, a reaction–transport framework, and evidence standards with an Au–Ag–Pd case study. *Catalysts* **2026**, *16*, 918. [CrossRef]
77. Kirsch, A.; Bøjesen, E.D.; Lefeld, N.; Larsen, R.; Mathiesen, J.K.; Skjærvø, S.L.; Pittkowski, R.K.; Sheptyakov, D.; Jensen, K.M.Ø. High-Entropy Oxides in the Mullite-Type Structure. *Chem. Mater.* **2023**, *35*, 8664–8674. [CrossRef] [PubMed]
78. Su, L.; Huyan, H.; Sarkar, A.; Gao, W.; Yan, X.; Addiego, C.; Kruk, R.; Hahn, H.; Pan, X. Direct observation of elemental fluctuation and oxygen octahedral distortion-dependent charge distribution in high entropy oxides. *Nat. Commun.* **2022**, *13*, 2358. [CrossRef] [PubMed]
79. Miao, L.; Sivak, J.T.; Kotsonis, G.; Ciston, J.; Ophus, C.L.; Dabo, I.; Maria, J.-P.; Sinnott, S.B.; Alem, N. Chemical Environment and Structural Variations in High Entropy Oxide Thin Film Probed with Electron Microscopy. *ACS Nano* **2024**, *18*, 14968–14977. [CrossRef] [PubMed]
80. He, M.; Davids, W.J.; Breen, A.J.; Ringer, S.P. Quantifying short-range order using atom probe tomography. *Nat. Mater.* **2024**, *23*, 1200–1207. [CrossRef] [PubMed]
81. Cieslak, J.; Reissner, M.; Berent, K.; Dabrowa, J.; Stygar, M.; Mozdzierz, M.; Zajusz, M. Magnetic properties and ionic distribution in high entropy spinels studied by Mössbauer and ab initio methods. *Acta Mater.* **2021**, *206*, 116600. [CrossRef]
82. Baek, J.; Hamkins, K.S.; Li, Y.; Garcia-Esparza, A.T.; Liu, T.; Kuo, C.-T.; Lee, J.-S.; Potter, A.W.; Kim, S.; Wang, Y.; et al. Modulating coordinate site occupancy in high-entropy spinel electrocatalysts. *Nat. Commun.* **2026**, *17*, 3540. [CrossRef] [PubMed]
83. Ueno, T.; Tsuji, Y. Discovery of Stable Short-Range Order in High-Entropy Oxides Assisted with Bayesian Optimization. *J. Chem. Theory Comput.* **2025**, *21*, 11697–11709. [CrossRef] [PubMed]
84. Chae, S.; Williams, L.; Lee, J.; Heron, J.T.; Kioupakis, E. Effects of local compositional and structural disorder on vacancy formation in entropy-stabilized oxides from first-principles. *npj Comput. Mater.* **2022**, *8*, 95. [CrossRef]
85. Ghigna, P.; Airoldi, L.; Fracchia, M.; Callegari, D.; Anselmi-Tamburini, U.; D'Angelo, P.; Pianta, N.; Ruffo, R.; Cibin, G.; de Souza, D.O.; et al. Lithiation Mechanism in High-Entropy Oxides as Anode Materials for Li-Ion Batteries: An Operando XAS Study. *ACS Appl. Mater. Interfaces* **2020**, *12*, 50344–50354. [CrossRef] [PubMed]
86. Tavani, F.; Fracchia, M.; Tofoni, A.; Braglia, L.; Jouve, A.; Morandi, S.; Manzoli, M.; Torelli, P.; Ghigna, P.; D'Angelo, P. Structural and mechanistic insights into low-temperature CO oxidation over a prototypical high entropy oxide by Cu L-edge operando soft X-ray absorption spectroscopy. *Phys. Chem. Chem. Phys.* **2021**, *23*, 26575–26584. [CrossRef] [PubMed]
87. De, S.; Bartók, A.P.; Csányi, G.; Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754–13769. [CrossRef] [PubMed]
88. Bartók, A.P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B* **2013**, *87*, 184115. Correction in *Phys. Rev. B* **2013**, *87*, 219902. Correction in *Phys. Rev. B* **2017**, *96*, 019902. [CrossRef]
89. Musil, F.; Grisafi, A.; Bartók, A.P.; Ortner, C.; Csányi, G.; Ceriotti, M. Physics-Inspired Structural Representations for Molecules and Materials. *Chem. Rev.* **2021**, *121*, 9759–9815. [CrossRef] [PubMed]
90. Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C.J.; Kornbluth, M.; Kozinsky, B. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **2023**, *14*, 579. [CrossRef] [PubMed]
91. Batatia, I.; Kovács, D.P.; Simm, G.; Ortner, C.; Csányi, G. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. In *Advances in Neural Information Processing Systems*; Curran Associates: Red Hook, NY, USA, 2022; Volume 35, pp. 11423–11436. [CrossRef]
92. Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J.P.; Kornbluth, M.; Molinari, N.; Smidt, T.E.; Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 2453. [CrossRef] [PubMed]
93. Jacobs, R.; Morgan, D.; Attarian, S.; Meng, J.; Shen, C.; Wu, Z.; Xie, C.Y.; Yang, J.H.; Artrith, N.; Blaiszik, B.; et al. A practical guide to machine learning interatomic potentials—Status and future. *Curr. Opin. Solid State Mater. Sci.* **2025**, *35*, 101214. [CrossRef]
94. da Silva Sousa Santos, A.; de Almeida, J.M.; de Jesus Trindade, F.; Felez, M.R.; da Silva Marques, L.; de Florio, D.Z. Machine learning assisted prediction and synthesis of new high-entropy fluorite oxide. *Open Ceram.* **2026**, *26*, 100982. [CrossRef]
95. Saal, J.E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65*, 1501–1509. [CrossRef]
96. Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N.E.R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; et al. Matminer: An open source toolkit for materials data mining. *Comput. Mater. Sci.* **2018**, *152*, 60–69. [CrossRef]

97. Leong, Z.; Desai, P.; Morley, N. Can Empirical Biplots Predict High Entropy Oxide Phases? *J. Compos. Sci.* **2021**, *5*, 311. [CrossRef]
98. da Silva Sousa Santos, A.; Stojanovska, E.; Alves, A.A.; de Paula, A.J.; de Florio, D.Z.; de Almeida, J.M. Rational Design of Single-Phase High-Entropy Oxides via Large Language Model Data Mining and Explainable Machine Learning. *J. Chem. Inf. Model.* **2026**, *66*, 5234–5248. [CrossRef] [PubMed]
99. Li, S.; Wei, S.; Huang, C.; Zhang, Y.; Zhang, G.; Sun, S. Extracting and reconstructing knowledge in materials science literature using large language models. *Commun. Mater.* **2026**, *7*, 31. [CrossRef]
100. Yang, F.; Chen, W.; Evans, J.D. Large language models in materials science and the need for open-source approaches. *Digit. Discov.* **2026**, *5*, 1981–1990. [CrossRef]
101. Deng, Y.; Li, Y.; Sai Gautam, G.; Zhu, B.; Deng, Z. Accelerating the discovery of disordered multi-component solid-state electrolytes using machine learning interatomic potentials. *J. Mater. Chem. A* **2025**, *13*, 34507–34518. [CrossRef]
102. Yuan, E.C.-Y.; Liu, Y.; Chen, J.; Zhong, P.; Raja, S.; Kreiman, T.; Vargas, S.; Xu, W.; Head-Gordon, M.; Yang, C.; et al. Foundation models for atomistic simulation of chemistry and materials. *Nat. Rev. Chem.* **2026**, *10*, 212–230. [CrossRef] [PubMed]
103. Kanno, M.; Taniike, T.; Honma, I. Data-Driven Exploration of Critical Factors for Single-Phase High-Entropy Oxide Anode Materials. *J. Phys. Chem. Lett.* **2025**, *16*, 9824–9829. [CrossRef] [PubMed]
104. Ma, B.; Dang, C.; Song, J.; Chen, Z.; Zhou, Y. Machine learning–accelerated discovery of novel high-entropy spinel oxide cathodes for solid oxide fuel cells. *Fuel* **2026**, *404*, 136190. [CrossRef]
105. Zhang, L.; Wan, Y.; Shibuta, Y.; Huang, X. Progress in machine learning interatomic potential and its applications in materials science. *Prog. Nat. Sci. Mater. Int.* **2025**, *35*, 1079–1104. [CrossRef]
106. 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. [CrossRef]
107. Liu, J.; Tang, Y.; Tretiak, S.; Duan, W.; Zhou, L. SA-GAT-SR: Self-adaptable graph attention networks with symbolic regression for high-fidelity material property prediction. *npj Comput. Mater.* **2025**, *11*, 377. [CrossRef]
108. Hoyos-Sanchez, J.P.; Hahn, H.; Jha, S.K.; Schweidler, S.; Velasco, L. A Machine Learning Approach to Determine the Band Gap Energy of High-Entropy Oxides Using UV-Vis Spectroscopy. *Eng* **2025**, *6*, 340. [CrossRef]
109. 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. [CrossRef] [PubMed]
110. Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S.P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31*, 3564–3572. [CrossRef]
111. Brahlek, M.; Gazda, M.; Keppens, V.; Mazza, A.R.; McCormack, S.J.; Mielewczyk-Gryń, A.; Musico, B.; Page, K.; Rost, C.M.; Sinnott, S.B.; et al. What is in a name: Defining "high entropy" oxides. *APL Mater.* **2022**, *10*, 110902. [CrossRef]
112. Otyepka, M.; Pykal, M.; Otyepka, M. Advancing materials discovery through artificial intelligence. *Appl. Mater. Today* **2025**, *47*, 102981. [CrossRef]
113. Goodall, R.E.A.; Lee, A.A. Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. *Nat. Commun.* **2020**, *11*, 6280. [CrossRef] [PubMed]
114. Jha, D.; Ward, L.; Paul, A.; Liao, W.-K.; Choudhary, A.; Wolverton, C.; Agrawal, A. ElemNet: Deep Learning the Chemistry of Materials from Only Elemental Composition. *Sci. Rep.* **2018**, *8*, 17593. [CrossRef] [PubMed]
115. Ward, L.; Agrawal, A.; Choudhary, A.; Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Comput. Mater.* **2016**, *2*, 16028. [CrossRef]
116. Potter, A.; Wang, Y.; Hamkins, K.; Kong, D.; Li, Y.; Qin, J.; Zheng, X. A statistical understanding of oxygen vacancies in distorted high-entropy perovskite oxides. *Nat. Commun.* **2026**, *17*, 4621. [CrossRef] [PubMed]
117. Zhang, M.; Duan, X.; Gao, Y.; Zhang, S.; Lu, X.; Luo, K.; Ye, J.; Wang, X.; Niu, Q.; Zhang, P.; et al. Tuning Oxygen Vacancies in Oxides by Configurational Entropy. *ACS Appl. Mater. Interfaces* **2023**, *15*, 45774–45789. [CrossRef] [PubMed]
118. Mancuso, A.; Monzillo, K.; Sacco, O.; Spiridigliozzi, L.; Monfreda, V.; Muscatello, A.; Dell'Agli, G.; Esposito, S.; Vaiano, V. Bixbyite-structured high-entropy oxide $(CeZrYbErGd)_2O_3+\delta$ for visible-light photocatalytic degradation of gallic acid: Performance and energy efficiency. *Catal. Today* **2026**, *473*, 115824. [CrossRef]
119. Mints, V.A.; Svane, K.L.; Rossmeisl, J.; Arenz, M. Exploring the High-Entropy Oxide Composition Space: Insights through Comparing Experimental with Theoretical Models for the Oxygen Evolution Reaction. *ACS Catal.* **2024**, *14*, 6936–6944. [CrossRef]
120. Xu, H.; Chen, T.; Shen, X.; Sun, N.; Wang, J.; Zheng, G.; Wang, X.; Ju, Y.; Xu, L.; Wang, S. A perovskite fuel electrode for efficient and sustainable CO_2 reduction in solid oxide electrolysis cells via A-site entropy engineering. *Fuel* **2026**, *417*, 138612. [CrossRef]
121. Butler, K.T.; Davies, D.W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555. [CrossRef] [PubMed]
122. Ramprasad, R.; Batra, R.; Pilania, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine learning in materials informatics: Recent applications and prospects. *npj Comput. Mater.* **2017**, *3*, 54. [CrossRef]
123. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

124. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
125. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
126. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]
127. Borshon, I.Z.; Ragone, M.; Phakatkar, A.H.; Long, L.; Shahbazian-Yassar, R.; Mashayek, F.; Yurkiv, V. Predicting column heights and elemental composition in experimental transmission electron microscopy images of high-entropy oxides using deep learning. *npj Comput. Mater.* **2024**, *10*, 275. [CrossRef]
128. Merchant, A.; Batzner, S.; Schoenholz, S.S.; Aykol, M.; Cheon, G.; Cubuk, E.D. Scaling deep learning for materials discovery. *Nature* **2023**, *624*, 80–85. [CrossRef] [PubMed]
129. Choudhary, K.; DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. *npj Comput. Mater.* **2021**, *7*, 185. Correction in *npj Comput. Mater.* **2021**, *8*, 221. [CrossRef]
130. Ma, Y.; Gao, Y.; Wang, L.; Chen, M.; Cui, W.; Wang, B.; Du, Y. Accelerating materials discovery through active learning: Methods, challenges and opportunities. *Innov. Inform.* **2025**, *1*, 100013. [CrossRef]
131. Dunn, A.; Wang, Q.; Ganose, A.; Dopp, D.; Jain, A. Benchmarking materials property prediction methods: The Matbench test set and Automatminer reference algorithm. *npj Comput. Mater.* **2020**, *6*, 138. Correction in *npj Comput. Mater.* **2020**, *6*, 159. [CrossRef]
132. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*; Curran Associates: Red Hook, NY, USA, 2017; Volume 30, pp. 4765–4774.
133. Cravero, F.; Ponzoni, I.; Diaz, M.F.; Vazquez, G.E. From prediction to understanding: A review of XAI applications and innovations in materials science. *Comput. Mater. Sci.* **2026**, *267*, 114493. [CrossRef]
134. Vassilev-Galindo, V.; Llorca, J. Explainable artificial intelligence for materials discovery: Application to catalysts for the HER and ORR. *Chem. Sci.* **2026**, *17*, 1058–1072. [CrossRef] [PubMed]
135. Zuo, D.; Tian, J.; Sun, Y.; Yu, X.; Peng, B.; Yu, T.; Xu, C.; Xu, X.; Wang, Y.; Liu, Y.; et al. Machine learning-accelerated discovery of multi-cation entropy-stabilized NASICON solid electrolytes with 10 000 hours of stable Na plating/stripping for all-solid-state sodium batteries. *Energy Environ. Sci.* **2026**, *19*, 860–869. [CrossRef]
136. Iwasaki, Y.; Toyama, R.; Tamura, R.; Matsuda, S.; Sakuraba, Y.; Kotsugi, M.; Igarashi, Y.; Sodeyama, K. Large-language-model-driven adaptive search space definition for autonomous closed-loop materials exploration. *Commun. Mater.* **2026**, *7*, 202. [CrossRef]
137. Katzer, B.; Klinder, S.; Schulz, K. Towards an automated workflow in materials science for combining multi-modal simulation and experimental information using data mining and large language models. *Mater. Today Commun.* **2025**, *45*, 112186. [CrossRef]
138. Miret, S.; Krishnan, N.M.A. Enabling large language models for real-world materials discovery. *Nat. Mach. Intell.* **2025**, *7*, 991–998. [CrossRef]
139. Ma, X.; Duan, X.; Qu, Y.; Gao, Y.; Wang, X.; Ru, K.; Zhang, M.; Niu, Q.; Zhang, P. Accelerating the discovery of high-entropy perovskite oxides by machine learning. *Chin. Chem. Lett.* **2026**, 112756. [CrossRef]
140. Yang, R.; Zhou, D.; Jia, Z.; Han, Y.; Tang, L.; Jiang, Z.; Tai, X.; Cai, Y.; Zhong, W.; Lin, Y.; et al. Toward Practical Design of High-Entropy Catalysts for Chlorine Evolution Reaction via Pareto-Guided Multi-Objective Bayesian Optimization Enabled by a Robotic AI-Chemist. *Angew. Chem. Int. Ed.* **2026**, *65*, e8794274. [CrossRef] [PubMed]
141. Niculescu, G.E.; Bejger, G.R.; Barber, J.P.; Wright, J.T.; Almishal, S.S.I.; Webb, M.; Ayyagari, S.V.G.; Maria, J.P.; Alem, N.; Heron, J.T.; et al. Local structure maturation in high entropy oxide $(Mg,Co,Ni,Cu,Zn)_{1-x}(Cr,Mn)xO$ thin films. *J. Am. Ceram. Soc.* **2025**, *108*, e20171. [CrossRef]
142. Olivetti, E.A.; Cole, J.M.; Kim, E.; Kononova, O.; Ceder, G.; Han, T.Y.-J.; Hiszpanski, A.M. Data-driven materials research enabled by natural language processing and information extraction. *Appl. Phys. Rev.* **2020**, *7*, 041317. [CrossRef]
143. Lin, C.-C.; Chang, C.-W.; Kaun, C.-C.; Su, Y.-H. Stepwise Evolution of Photocatalytic Spinel-Structured $(Co,Cr,Fe,Mn,Ni)_3O_4$ High Entropy Oxides from First-Principles Calculations to Machine Learning. *Crystals* **2021**, *11*, 1035. [CrossRef]
144. Elmutasim, O.; Hussien, A.G.; Sharan, A.; AlKhoori, S.; Vasiliades, M.A.; Taha, I.M.A.; Kim, S.; Harfouche, M.; Emwas, A.-H.; Anjum, D.H.; et al. Evolution of Oxygen Vacancy Sites in Ceria-Based High-Entropy Oxides and Their Role in N_2 Activation. *ACS Appl. Mater. Interfaces* **2024**, *16*, 23038–23053. [CrossRef] [PubMed]
145. Settles, B. *Active Learning Literature Survey*; Computer Sciences Technical Report 1648; University of Wisconsin–Madison: Madison, WI, USA, 2009.
146. Batchelor, T.A.A.; Löffler, T.; Xiao, B.; Krysiak, O.A.; Strotkötter, V.; Pedersen, J.K.; Clausen, C.M.; Savan, A.; Li, Y.; Schuhmann, W.; et al. Complex-Solid-Solution Electrocatalyst Discovery by Computational Prediction and High-Throughput Experimentation. *Angew. Chem. Int. Ed.* **2021**, *60*, 6932–6937. [CrossRef] [PubMed]
147. Wu, L.-T.; Li, Z.-L.; Yen, S.-Y.; Kaghazchi, P.; Jiang, J.-C. Probing entropic control of stacking phase preference in layered oxide cathodes for sodium-ion batteries via machine-learning potentials. *npj Comput. Mater.* **2026**, *12*, 86. [CrossRef]

148. Riley, C.; De La Riva, A.; Park, J.E.; Percival, S.J.; Benavidez, A.; Coker, E.N.; Aidun, R.E.; Paisley, E.A.; Datye, A.; Chou, S.S. A High Entropy Oxide Designed to Catalyze CO Oxidation Without Precious Metals. *ACS Appl. Mater. Interfaces* **2021**, *13*, 8120–8128. [CrossRef] [PubMed]
149. Fan, M.Y.; Wang, J.J.; Zhao, J.; Zhang, H.; Ma, T.Y.; Han, X.P.; Hu, W.B. High-entropy oxide-supported platinum nanoparticles for efficient hydrogen evolution reaction. *Rare Met.* **2024**, *43*, 1537–1546. [CrossRef]
150. Svane, K.L. Prospects of using high entropy oxides as catalysts for the oxygen evolution reaction. *Curr. Opin. Electrochem.* **2025**, *51*, 101670. [CrossRef]
151. Gautam, A.; Ahmad, M.I. Li and Na substituted (Fe,Co,Ni,Cu,Zn)O high entropy oxide catalysts for oxygen evolution reaction. *Catal. Sci. Technol.* **2026**, *16*, 1995–2008. [CrossRef]
152. Kante, M.V.; Weber, M.L.; Ni, S.; van den Bosch, I.C.G.; van der Minne, E.; Heymann, L.; Falling, L.J.; Gauquelin, N.; Tsvetanova, M.; Cunha, D.M.; et al. A High-Entropy Oxide as High-Activity Electrocatalyst for Water Oxidation. *ACS Nano* **2023**, *17*, 5329–5339. [CrossRef] [PubMed]
153. Shan, B.-F.; Yang, J.; Xiang, X.; Zhao, Z.-Y. High-entropy compounds for photo(electro)catalysis: Diverse materials and applications. *J. Mater. Chem. A* **2025**, *13*, 12808–12827. [CrossRef]
154. Pedersen, J.K.; Batchelor, T.A.A.; Bagger, A.; Rossmeisl, J. High-Entropy Alloys as Catalysts for the CO_2 and CO Reduction Reactions. *ACS Catal.* **2020**, *10*, 2169–2176. [CrossRef]
155. Batchelor, T.A.A.; Pedersen, J.K.; Winther, S.H.; Castelli, I.E.; Jacobsen, K.W.; Rossmeisl, J. High-Entropy Alloys as a Discovery Platform for Electrocatalysis. *Joule* **2019**, *3*, 834–845. [CrossRef]
156. Saidi, W.A. Emergence of local scaling relations in adsorption energies on high-entropy alloys. *npj Comput. Mater.* **2022**, *8*, 86. [CrossRef]
157. Svane, K.L.; Rossmeisl, J. Theoretical Optimization of Compositions of High-Entropy Oxides for the Oxygen Evolution Reaction. *Angew. Chem. Int. Ed.* **2022**, *61*, e202201146. [CrossRef] [PubMed]
158. Baek, J.; Hossain, M.D.; Mukherjee, P.; Lee, J.; Winther, K.T.; Leem, J.; Jiang, Y.; Chueh, W.C.; Bajdich, M.; Zheng, X. Synergistic effects of mixing and strain in high entropy spinel oxides for oxygen evolution reaction. *Nat. Commun.* **2023**, *14*, 5936. [CrossRef] [PubMed]
159. Song, H.; Glezakou, V.-A.; Wu, Z.; Jiang, D.-E. Computational insights into hydrogen adsorption energies on medium-entropy oxides. *Catal. Sci. Technol.* **2025**, *15*, 4937–4944. [CrossRef]
160. Bligaard, T.; Nørskov, J.K.; Dahl, S.; Matthiesen, J.; Christensen, C.H.; Sehested, J. The Brønsted–Evans–Polanyi relation and the volcano curve in heterogeneous catalysis. *J. Catal.* **2004**, *224*, 206–217. [CrossRef]
161. Vojvodic, A.; Calle-Vallejo, F.; Guo, W.; Wang, S.; Toftelund, A.; Studt, F.; Martínez, J.I.; Shen, J.; Man, I.C.; Rossmeisl, J.; et al. On the behavior of Brønsted-Evans-Polanyi relations for transition metal oxides. *J. Chem. Phys.* **2011**, *134*, 244509. [CrossRef] [PubMed]
162. Fracchia, M.; Ghigna, P.; Pozzi, T.; Anselmi Tamburini, U.; Colombo, V.; Braglia, L.; Torelli, P. Stabilization by Configurational Entropy of the Cu(II) Active Site during CO Oxidation on Mg0.2Co0.2Ni0.2Cu0.2Zn0.2O. *J. Phys. Chem. Lett.* **2020**, *11*, 3589–3593. [CrossRef] [PubMed]
163. Wali, N.; Sabir, M.; Bilal, M.; Qadri, A.N.; Khan, A.; Yao, G.; Ji, Y.; Ullah, S.; Ur Rehman, N. From plasma-generated radicals to value-added products: A critical review of methane valorisation. *Catalysts* **2026**, *16*, 777. [CrossRef]
164. Edalati, P.; Wang, Q.; Razavi-Khosroshahi, H.; Fuji, M.; Ishihara, T.; Edalati, K. Photocatalytic hydrogen evolution on a high-entropy oxide. *J. Mater. Chem. A* **2020**, *8*, 3814–3821. [CrossRef]
165. Chen, L.; Fan, K.; Yan, L.; Yang, J. Photocatalytic reduction of CO_2 using a high-entropy fluorite oxide. *Ceram. Int.* **2025**, *51*, 50621–50628. [CrossRef]
166. Wang, S.-Y.; Chen, T.-Y.; Kuo, C.-H.; Lin, C.-C.; Huang, S.-C.; Lin, M.-H.; Wang, C.-C.; Chen, H.-Y. Operando synchrotron transmission X-ray microscopy study on (Mg, Co, Ni, Cu, Zn)O high-entropy oxide anodes for lithium-ion batteries. *Mater. Chem. Phys.* **2021**, *274*, 125105. [CrossRef]
167. Nguyen, T.T.; Arita, M.; Saeki, R.; Edalati, K. High-entropy oxide Ti-Zr-Hf-Nb-Ta-O as photoanode for photoelectrocatalytic hydrogen generation. *Scr. Mater.* **2025**, *264*, 116698. [CrossRef]
168. Nguyen, T.T.; Hidalgo-Jiménez, J.; Edalati, K. High-entropy oxide photocatalysts for green ammonia synthesis from nitrogen fixation in water. *Scr. Mater.* **2025**, *269*, 116931. [CrossRef]
169. Ling, H.; Sun, H.; Lu, L.; Zhang, J.; Liao, L.; Wang, J.; Zhang, X.; Lan, Y.; Li, R.; Lu, W.; et al. Sustainable photocatalytic hydrogen peroxide production over octonary high-entropy oxide. *Nat. Commun.* **2024**, *15*, 9505. [CrossRef] [PubMed]
170. Li, Y.; Chen, Z.; Liu, J.; Liu, R.; Zhang, C.; Li, H. Novel high entropy oxide as anode for high performance lithium-ion capacitors. *Ceram. Int.* **2023**, *49*, 38439–38447. [CrossRef]
171. Wang, K.; Hua, W.; Huang, X.; Stenzel, D.; Wang, J.; Ding, Z.; Cui, Y.; Wang, Q.; Ehrenberg, H.; Breitung, B.; et al. Synergy of cations in high entropy oxide lithium ion battery anode. *Nat. Commun.* **2023**, *14*, 1487. [CrossRef] [PubMed]

172. Ragul, S.; Kannadasan, K.; Elumalai, P. High-stability and high-capacity aqueous sodium-ion battery using a high-entropy oxide cathode: Intercalation vs. capacitive sodium charge storage. *Sustain. Energy Fuels* **2026**, *10*, 2256–2270. [CrossRef]
173. Zhou, Y.; Yu, Z.; Ouyang, B.; Yuan, Y.; Wang, Y.; Qin, F.; Lei, C.; Yang, Y.; Liu, W.; Ning, F.; et al. Tuning crystal field stabilization energy through configurational entropy design in high-entropy oxide sodium-ion battery cathodes. *Mater. Today* **2026**, *93*, 103192. [CrossRef]
174. Zhang, P.; Liu, J.; Zhang, T.; Zhou, X.; Shen, X.; Zhou, J.; Wang, S.; Zhang, L.; Wang, J.; Wang, M.; et al. Achieving High-Voltage Sodium–Ion Batteries with Phase-Transition-Free High-Entropy Oxide Cathodes via Altering Local Coordination Fields. *ACS Nano* **2026**, *20*, 10012–10023. [CrossRef] [PubMed]
175. Cai, H.; Zhang, P.; Li, B.; Zhu, Y.; Zhang, Z.; Guo, W. High-entropy oxides for energy-related electrocatalysis. *Mater. Today Catal.* **2024**, *4*, 100039. [CrossRef]
176. Wang, L.; You, Y.; Li, Z.; Ou, S.; Xu, J.; Yang, N.; Luo, H.; Yuan, M. Achieving highly reversible anionic redox and negligible voltage decay by a high-entropy strategy in Na-ion layered oxide cathodes. *J. Colloid Interface Sci.* **2026**, *701*, 138686. [CrossRef] [PubMed]
177. Hidalgo-Jiménez, J.; Akbay, T.; Ishihara, T.; Edalati, K. Investigation of a high-entropy oxide photocatalyst for hydrogen generation by first-principles calculations coupled with experiments: Significance of electronegativity. *Scr. Mater.* **2024**, *250*, 116205. [CrossRef]
178. Kusne, A.G.; Yu, H.; Wu, C.; Zhang, H.; Hattrick-Simpers, J.; DeCost, B.; Sarker, S.; Oses, C.; Toher, C.; Curtarolo, S.; et al. On-the-fly closed-loop materials discovery via Bayesian active learning. *Nat. Commun.* **2020**, *11*, 5966. [CrossRef] [PubMed]
179. Cheng, M.; Fu, C.-L.; Okabe, R.; Chotrattanapituk, A.; Boonkird, A.; Hung, N.T.; Li, M. Artificial intelligence-driven approaches for materials design and discovery. *Nat. Mater.* **2026**, *25*, 174–190. [CrossRef] [PubMed]
180. Zeni, C.; Pinsler, R.; Zügner, D.; Fowler, A.; Horton, M.; Fu, X.; Wang, Z.; Shysheya, A.; Crabbé, J.; Ueda, S.; et al. A generative model for inorganic materials design. *Nature* **2025**, *639*, 624–632. [CrossRef] [PubMed]
181. Jiang, X.; Wang, W.; Tian, S.; Wang, H.; Lookman, T.; Su, Y. Applications of natural language processing and large language models in materials discovery. *npj Comput. Mater.* **2025**, *11*, 79. [CrossRef]
182. Cole, J.M. A Design-to-Device Pipeline for Data-Driven Materials Discovery. *Acc. Chem. Res.* **2020**, *53*, 599–610. [CrossRef] [PubMed]
183. Volk, A.A.; Abolhasani, M. Performance metrics to unleash the power of self-driving labs in chemistry and materials science. *Nat. Commun.* **2024**, *15*, 1378. [CrossRef] [PubMed]
184. Yoon, Y.; Gu, G.H.; Kim, K. Closed-loop workflow of high-entropy materials discovery: Efficient and accurate synthesizability prediction via domain-specific local LLMs. *npj Comput. Mater.* **2026**, *12*, 221. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.