

Interpretable data-driven discovery of high-performance multielement metal oxide electrocatalysts for oxygen evolution reaction from a small hybrid dataset

Wenqin Peng, Shigenobu Hayashi, Abraham Castro Garcia, Lei Fang, Masaki Takeguchi, Yibin Xu & Ken Sakaushi

To cite this article: Wenqin Peng, Shigenobu Hayashi, Abraham Castro Garcia, Lei Fang, Masaki Takeguchi, Yibin Xu & Ken Sakaushi (2026) Interpretable data-driven discovery of high-performance multielement metal oxide electrocatalysts for oxygen evolution reaction from a small hybrid dataset, Science and Technology of Advanced Materials, 27:1, 2713992, DOI: 10.1080/14686996.2026.2713992

To link to this article: <https://doi.org/10.1080/14686996.2026.2713992>



© 2026 The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group.



View supplementary material [↗](#)



Published online: 20 Aug 2026.



Submit your article to this journal [↗](#)



Article views: 136



View related articles [↗](#)



View Crossmark data [↗](#)

Interpretable data-driven discovery of high-performance multielement metal oxide electrocatalysts for oxygen evolution reaction from a small hybrid dataset

Wenqin Peng^a, Shigenobu Hayashi^b, Abraham Castro Garcia^a, Lei Fang^b, Masaki Takeguchi^a, Yibin Xu^b and Ken Sakaushi^{a,c,d}

^aResearch Center for Energy and Environmental Materials, National Institute for Materials Science, Tsukuba, Japan;

^bCenter for Basic Research on Materials, National Institute for Materials Science, Tsukuba, Japan;

^cChurchill College, Cambridge, UK;

^dCavendish Laboratory, University of Cambridge, Cambridge, UK

ABSTRACT

One of the most formidable challenges in materials chemistry is the rational design of functionalities capable of dramatically enhancing performance. However, it is well-known that the discovery of promising materials often requires several decades of continuous trial-and-error. Herein, we show an interpretable data-driven framework for the discovery of multielement metal oxide oxygen evolution reaction (OER) electrocatalysts in alkaline media within a substantially shorter timeframe. This framework was trained on a hybrid dataset comprising only 557 data, consisting of the curated literature and our own experimental results. Furthermore, this framework was specifically designed to enable extrapolative materials discovery, including the exploration of elemental combinations absent from the training database. Consequently, from a large material search space of approximately 3 million candidates, we identified a promising unconventional quinary oxide composed of V, Ni, W, Rh, and Ru that exhibits, in 0.1 M KOH, an exchange current density approximately 20 times higher than that of IrO₂. This work serves as a proof-of-concept, demonstrating that the rational design of high-performance electrochemical functionalities from an extensive candidate space can be achieved using a small hybrid dataset combined with an interpretable data-driven approach.

ARTICLE HISTORY

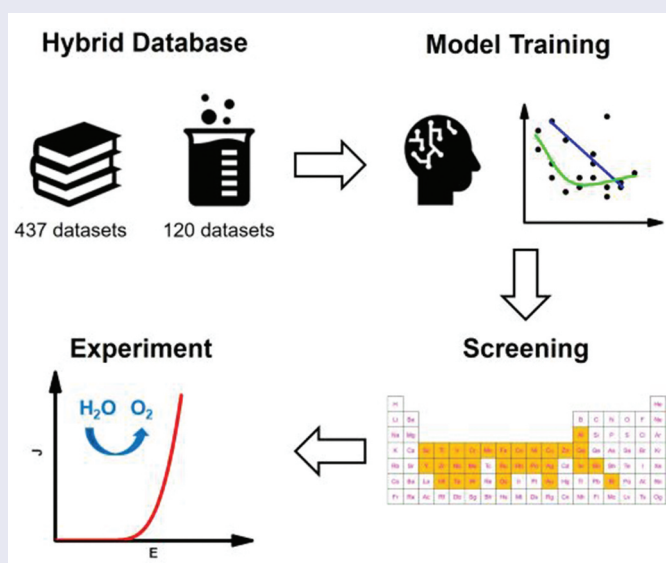
Received 24 June 2026

Revised 19 July 2026

Accepted 30 July 2026

KEYWORDS

Data-driven material discovery; small data; electrocatalysts; interpretable data-driven approach; multielement metal oxide; oxygen evolution reaction



IMPACT STATEMENT

This work develops an extrapolative-sniping artificial intelligence framework based on a small hybrid database comprising 557 data, accelerating the discovery of multielement metal oxide electrocatalysts with excellent oxygen evolution performance.

CONTACT Ken Sakaushi  sakaushi.ken@nims.go.jp  Research Center for Energy and Environmental Materials, National Institute for Materials Science, Namiki 1-1, Tsukuba 305-0044, Japan

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/14686996.2026.2713992>

© 2026 The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Introduction

Finding a high-performance functionality with affordable element demands an extraordinary effort even for a single technology: Paul Alwin Mittasch, a co-worker of Fritz Haber, had attempted several tens of thousands of systematic experiments to check multicomponent material candidates combining 54 elements towards the development of the Haber-Bosch process [1,2]. The motivation for this hectic mission is not only for discovering a cost-effective catalyst with better activity, but also to understand why the specific iron oxide from Gällivare, Sweden, is able to act as a good catalyst for this process but not iron oxide from other regions. Thus, systematic experiments with large material search space are indispensable to find advanced materials towards developing modern technology and to unveil basic mechanisms of complicated chemical reactions or building up design principles to synthesize materials with desirable functions. From this background, it is of great interest to establish effective methodology equipped with artificial intelligence (AI) to shorten the trial-and-error timeframe during systematic experiments to find a material with target properties [3–6]. In particular, this methodology is important for the development of advanced electrochemical devices such as water electrolyzers, a critical climate technology for the green hydrogen production. This is because of the urgent demand for high-performance electrocatalysts for oxygen evolution reaction (OER) to meet carbon neutrality goals by 2050 [7], however discovery of promising electrochemical materials often requires several decades of continuous efforts due to complexity of electrode processes [8–11]. Moreover, multielement materials are attracting great interest due to their enhanced properties originating from non-trivial element combining effects [12–16]. However, optimization of this type of advanced materials is even more challenging compared to traditional materials. Aiming to develop high-performance multielement OER electrocatalysts, Robotics-AI combination is known to be capable of dealing with a larger search space containing millions or more of candidates, as demonstrated by

its overwhelming power in several fields of materials discovery, for instance drug discovery [17–21]. However in the field of electrochemical material discovery, the unique features of electrochemical experiments cause the robotic experiments to require a long and expensive period to establish an optimized protocol to provide accurate and reliable data [22]. Moreover, a promising alternative approach of Human-AI collaboration has been proven to work however only in a candidate search space with several thousands of candidates so far [22,23]. From this point, a material prediction AI is an alternative to robotics-AI or human-AI approaches for a material discovery from a large search space.

Here, we have developed an interpretable data-driven framework for the discovery of high-performance multielement metal oxide electrocatalysts toward the alkaline oxygen evolution reaction from a large material search space using a small hybrid dataset. A data-driven approach with interpretability and/or explainability is essential for ‘opening’ black-boxes of present standard AI-driven material discovery [24–26]. Our approach stands on the shoulders of past and present researchers: we curated only 437 data points of transition metal oxide-based oxygen evolution electrocatalysts from the literature and integrated these curated datasets with 120 datasets derived from our own experiments to construct a hybrid database. This small hybrid database was employed to discover multielement materials using an extrapolative material prediction algorithm with 35 parameters, navigating a large search space of approximately 3 million candidates, including elements not presented in our original database. As a result, we identified a promising oxide composed of metal elements of V, Ni, W, Rh, and Ru that delivers excellent oxygen evolution performance, which was validated through experimental measurements. This work serves as a proof-of-concept, demonstrating that a promising material discovery from a large search space is able to be achieved by using a small hybrid data with datasets in the order of 10^2 , combined with a highly accurate and reliable algorithm with 35 parameters (Figure 1).

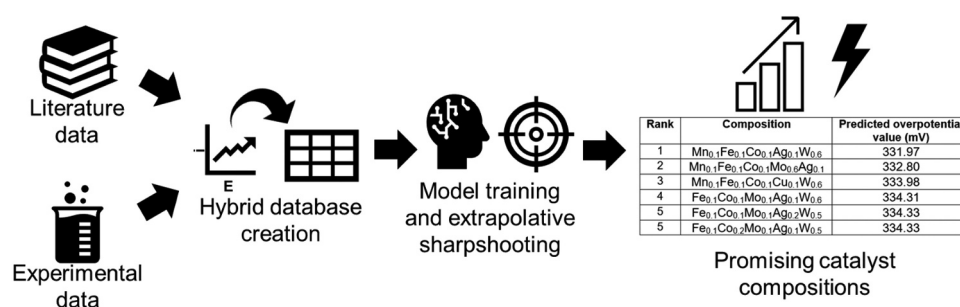


Figure 1. Overview of applied data-driven electrocatalyst search approach.

Results and discussion

To develop the hybrid database, papers from the existing literature were collected based on expert-judgement. Our search for relevant literature resulted in 203 papers, which were then systematically analyzed, extracting the data describing experimental conditions and electrocatalyst composition. This was done by digitizing or extrapolating from graphs as necessary to match overpotential at $1 \text{ mA/cm}^2_{\text{geo}}$ with electrocatalyst composition and electrolyte pH, ensuring that all data is on a standardized scale for fair comparison. This resulted in a literature-based data set consisting of 437 data points comprised entirely of oxide-based OER electrocatalysts. The composition of the electrocatalysts was represented using 17 group coefficients (belonging to groups 1 to 17 of the periodic table) and 17 group weighted properties, for a total of 35 features, including electrolyte pH [27]. The calculation of these group coefficients and group weighted properties is explained in detail in the Experimental Section. Having finished curating this literature data set an additional 120 datapoints from our own previous experiments controlled under the identical conditions were added to create a hybrid database. The literature dataset covered a broad range of reported compositions and included many high-performance catalysts, whereas the systematically acquired in-house dataset supplied internally consistent data and a substantial number of relatively high-overpotential materials that are underrepresented in the literature. The two datasets provided complementary information since the systematic experimental dataset partially compensates for the heterogeneity and publication bias of the literature dataset, while the literature dataset expands the compositional coverage beyond that of the in-house experiments.

Machine learning models: Gaussian Process regression, with radial basis function kernel (GPR (rbf)), Matern kernel (GPR (mt)) and rational quadratic kernel (GPR (rq)), Natural Gradient Boosting (NGB), Random Forest (RF) and Extreme Gradient Boosting (XGB) were trained on the literature dataset, experimental dataset and finally the combined, hybrid

database. The performance of the models shows that the combination of both literature and data from our own experiments yielded the best model performance as can be observed in Table 1.

To investigate why using the combined dataset resulted in significantly improved model performance, the various data sets; experimental, literature and their combination were analyzed. In Figure 2(a), it can be observed that for Groups 4, 5, 6, 10, 11 and 12, experimental data constitutes most of the data points present in the combined data set. In Figure 2(b), the distribution of overpotentials observed in the experimental data set indicates the presence of many high overpotential electrocatalysts (600–1000 mV), in contrast to the distribution from the literature data set observed in Figure 2(c), which shows an abundance of low overpotential electrocatalysts centered around 200–300 mV. Figure 2(d) shows the distribution that results from the combination of experimental and literature datasets. This combined data set is more robust, containing various examples of high and low overpotential catalysts with diverse elemental compositions. This matches well with both older [28,29] and newer [30] perspectives, pointing out that quality and diversity in data is important. The presence of ‘bad’ electrocatalysts in the dataset is crucial to enhance the likelihood of efficiently finding new promising electrocatalyst compositions.

Of the ML models trained XGB showed the best performance, although it is generally considered to be a ‘black box’ due to its lack of immediate interpretability, we decided to compute SHAP (SHapley Additive exPlanations) values to better understand the logic through which the model predicts overpotential of a given electrocatalyst as a function of the features available in the dataset. SHAP values can be used as a method to interpret ML models, although originally come from cooperative game theory. A more detailed explanation of the application of SHAP values in ML is available from Benedek, et al.[31]

In Figure 3(a), the SHAP bar plot shows that only a few of the 35 available features in the data have high importance, notably, Group coefficients 10, 9, 8, 6 and

Table 1. Accuracy (R^2) of ML models for overpotential @ $1 \text{ mA/cm}^2_{\text{geo}}$.

	GPR (rbf)	GPR (rq)	NGB	RF	XGB
Using experimental data (120 data points)					
Train	0.865	0.885	0.996	0.963	0.934
Test	0.635	0.632	0.61	0.579	0.621
Using literature data (437 data points)					
Train	0.668	0.924	0.78	0.885	0.885
Test	0.437	0.453	0.579	0.595	0.581
Using hybrid database (557 data points)					
Train	0.946	0.967	0.971	0.982	0.949
Test	0.822	0.819	0.853	0.84	0.877

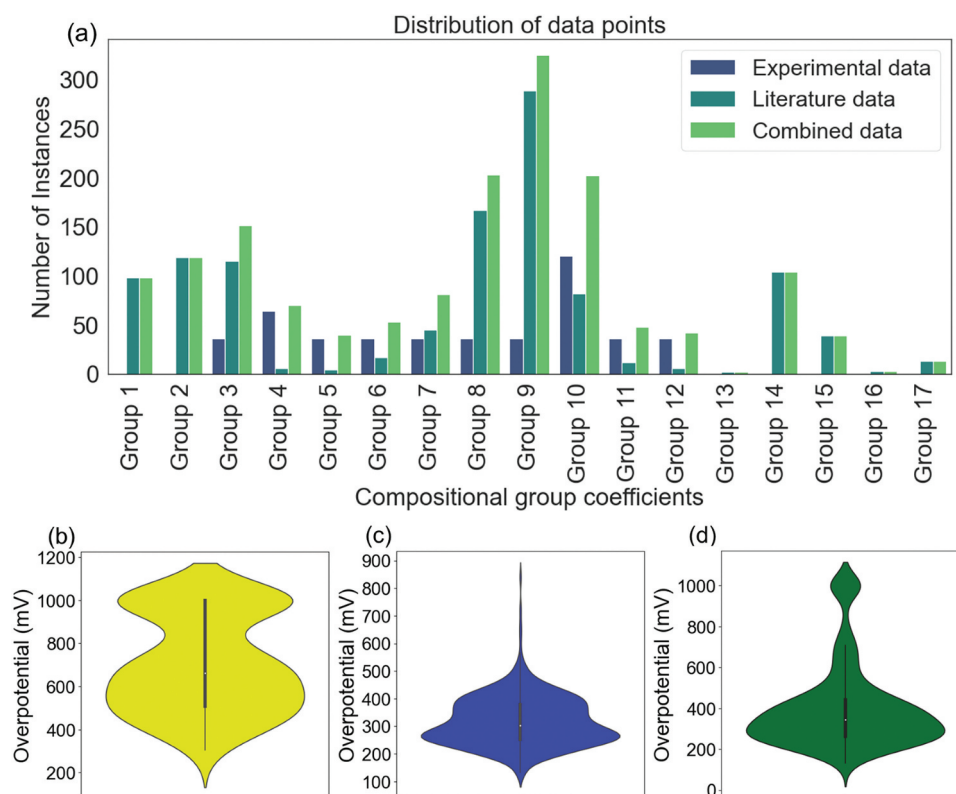


Figure 2. Analysis of data distribution in experimental, literature and combined data sets. (a) Bar plot comparatively showing the number of data points in each respective data set for all compositional group coefficients. (b)–(d) Violin plots showing the distribution of data points for experimental, literature and combined data sets.

pH (nominal) hold high importance. The aggregate bottom 28 features have a comparatively smaller importance, indicating that the predictions depend mostly on a few groups of elements. In Figure 3(b), the SHAP bee swarm plot, where we can observe that the absence (blue) or presence (purple to red) of certain elements is correlated with electrocatalyst overpotential. For group coefficient 10, a low value (blue, indicating the absence of elements from this group) seems to be correlated with lower overpotential, meanwhile group coefficients 9, 8 and 6 show that higher values (purple, red) show a correlation with low overpotential, indicating that their presence improves the performance of the electrocatalyst.

Previous explainable-AI studies have shown that SHAP analysis is effective for identifying composition–activity relationships in predefined multielement OER catalyst systems. For example, SHAP has been used to quantify the individual and interactive effects of Ni, Fe, Co, and Ce in (Ni–Fe–Co–Ce) O_x catalysts, and more recently to interpret the optimization of elemental ratios in Fe–Co–Cr–Mn–Cu high-entropy layered double hydroxides [32,33]. In these studies, the elemental candidates were fixed and the principal objective was to optimize their relative compositions. In contrast, the present model represents a broader range of oxide catalysts using periodic-table group descriptors. The resulting SHAP analysis therefore identifies compositional trends across different

catalyst families rather than within a single predefined chemical system. The importance of Groups 8, 9, and 10 is broadly consistent with the central role of late transition metals frequently observed in active OER catalysts, while the identified contribution of Group 6 suggests that elements from this group may provide an additional compositional degree of freedom in multi-element catalyst design. The importance of pH further demonstrates that the model captures not only catalyst composition but also the influence of the electrochemical environment.

Based on the satisfactory ML model performance achieved with the hybrid dataset and our understanding of it, an extrapolative screening approach was used to suggest new quinary oxide electrocatalyst compositions, looking to minimize overpotential at 1 mA/cm²_{geo}. This approach considered 28 elements of the periodic table and assumed that the catalyst formula would be $E_1x_1E_2x_2E_3x_3E_4x_4E_5x_5$, where E stands for a given element and x_i can take values of 0.5 or 1. This resulted in a potential experimental space consisting of 3.04 million possible electrocatalyst compositions. The 5 trained ML models were used to calculate the overpotential @ 1 mA/cm²_{geo} for all 3.04 million candidates, of which the top 3000 candidates with the lowest calculated overpotential were tabulated and ranked as follows. Primary ranking: rank the candidates by the number of models recommending it, a larger number corresponding to a higher rank and secondary ranking:

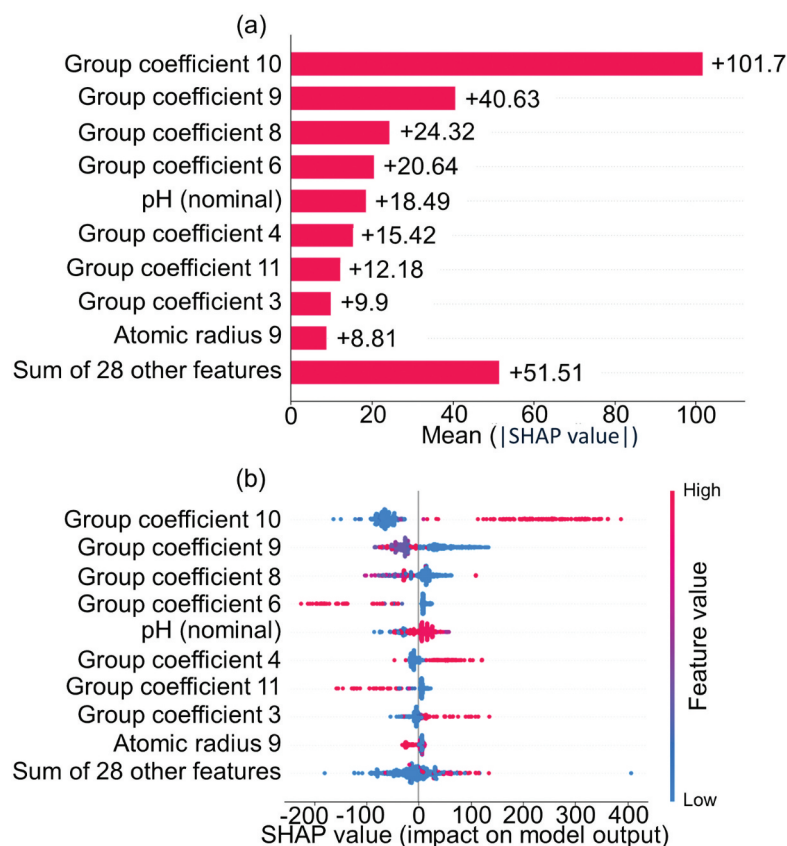


Figure 3. Different representations of SHAP values computed for XGB model trained using the combined dataset. (a) SHAP bar plot showing the importance of most important features. (b) SHAP bee swarm plot displaying the distribution of impact of each feature.

for candidates recommended by 2 models, further rank them by the average rank of the two models. The presumption behind this approach is that we do not trust any model more than others, given that they all achieved acceptable predictive performance. Subsequently, the most promising candidates calculated and ranked were synthesized and tested.

The five quinary-metal oxide-based compounds (listed in Table S1), which were predicted as top-ranked materials by our AI-based approach, were synthesized using a dip-coating method. The OER activities of the five materials were investigated in a three-electrode two-compartment cell configuration in the 0.1 M KOH electrolyte (pH13). We note here that the mass-loading for all the electrocatalysts is 0.13 mg. Figure 4(a) shows the iR -corrected linear sweeping voltammetry (LSV) curves for OER. The materials with metal element composition ratio of Ru:Rh:V:Ni:W = 0.5:0.5:0.5:0.5:1 (the $\text{Ru}_{0.5}\text{Rh}_{0.5}\text{V}_{0.5}\text{Ni}_{0.5}\text{W}$ oxide; denoted as EL-4) were found to exhibit the best OER activity among these five electrocatalysts, achieving a current density of 81.1 mA/cm^2 at overpotential of 0.5 V, higher than that of the in-house-prepared benchmark IrO_2 (48.3 mA/cm^2 at overpotential of 0.5 V). Furthermore, EL-4 shows the lowest overpotentials of 201.5 mV at a current density of 1 mA/cm^2 and of 315.2 mV at 10 mA/cm^2 (Figure 4(b)). The

overpotential values at 10 mA/cm^2 are 378.6 mV and 362.3 mV for EL-2 ($\text{Rh}_{0.5}\text{Os}_{0.5}\text{V}_{0.5}\text{Ni}_{0.5}\text{W}$) and EL-3 ($\text{Ru}_{0.5}\text{Rh}_{0.5}\text{Ni}_{0.5}\text{Nb}_{0.5}\text{W}$), respectively, which are comparable to that of IrO_2 (375.4 mV), while it increases to 408.8 mV for EL-1 ($\text{Rh}_{0.5}\text{Os}_{0.5}\text{Ni}_{0.5}\text{Nb}_{0.5}\text{W}$). However, EL-5 ($\text{Ru}_{0.5}\text{Pd}_{0.5}\text{Os}_{0.5}\text{Nb}_{0.5}\text{W}$), in which Rh is absent, has a relatively higher overpotential of 441.8 mV at 10 mA/cm^2 .

As for a multistep electrochemical reaction of OER, the Tafel slope suggests information about reaction kinetics. Here the Tafel slopes were derived from their Tafel plots shown in Figure 4(c). EL-4 exhibits a smaller Tafel slope of 92.6 mV/dec, with respect to EL-1 (93.5 mV/dec), EL-3 (98.2 mV/dec) and EL-5 (126.6 mV/dec), suggesting a similar kinetics in OER except EL-5. Furthermore, mass activity (normalized by the mass loading of the electrocatalyst) is used to reveal intrinsic activity of electrocatalysts (Figure S1). At an overpotential of 300 mV, the mass activity of EL-4 is 30.3 A/g, which is more than 2 fold higher than that of EL-1 (2.5 A/g), EL-2 (5.2 A/g), EL-3 (12.3 A/g) and EL-5 (2.9 A/g). The chronopotentiometric measurement was carried out to test the durability of these electrocatalysts at a constant current density of 50 mA/cm^2 for 3 h. As shown in Figure S2, these materials exhibit comparable durability with respect to IrO_2 in alkaline electrolyte. Additionally, the long-term

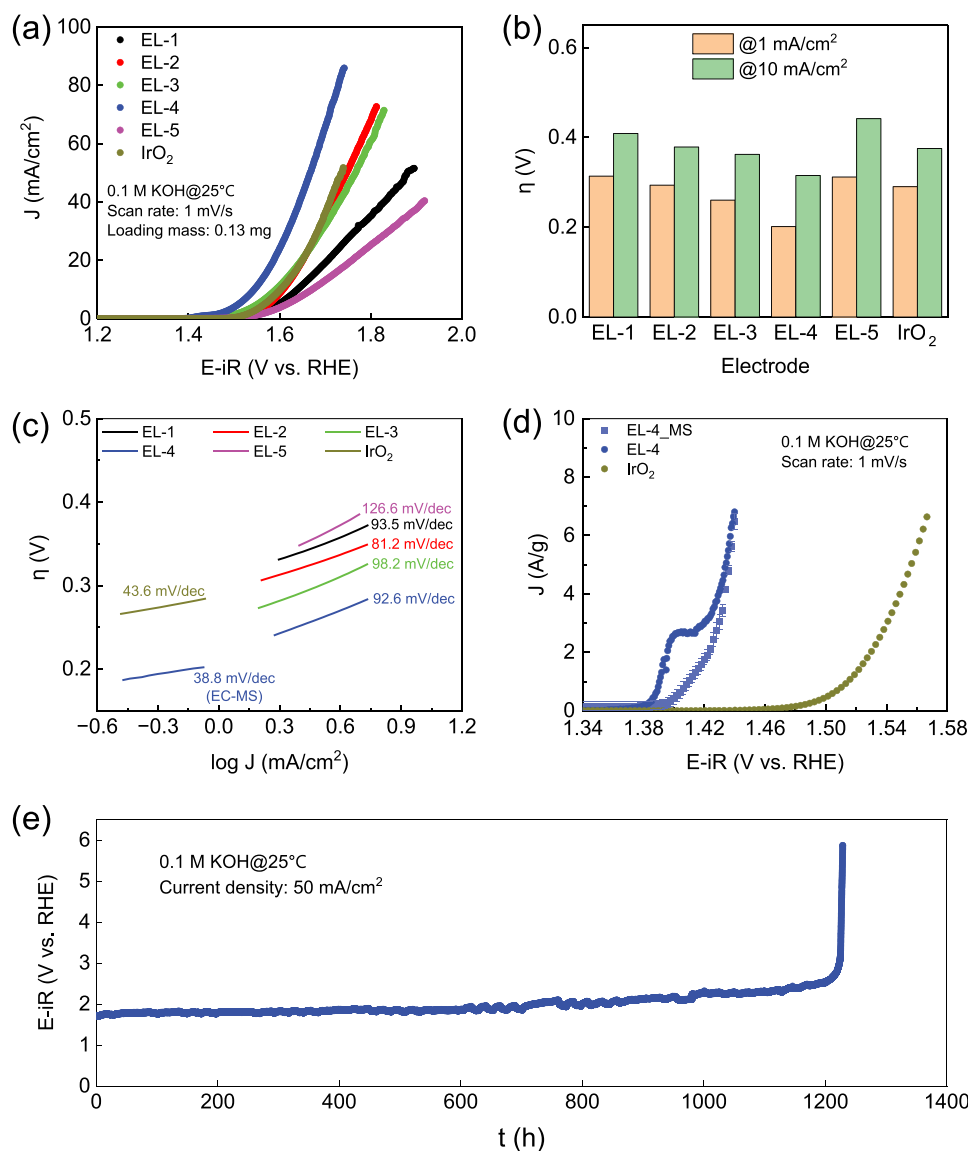


Figure 4. OER performance of quinary-metal-based electrocatalysts measured in 0.1 M KOH electrolyte. (a) LSV curves measured at a scan rate of 1 mV/s. (b) Overpotentials at current densities of 1 and 10 mA/cm². (c) Corresponding Tafel plots. (d) Specific mass activities of EL-4 and IrO₂ at the low potential region. Due to the redox peak, a real-time monitoring electrochemical mass-spectroscopy (EC-MS) was employed to obtain the data (square: OER oxygen current based on EC-MS measurement of EL-4). (e) Long-term durability test of EL-4 at a current density of 50 mA/cm².

durability of EL-4 was further investigated. At the current density of 50 mA/cm², EL-4 exhibits stable water oxidation operation over 1000 h (Figure 4(e)), suggesting the promising properties of EL-4. The activity at low potentials, e.g. regions close to onset potentials, is especially crucial for evaluating electrocatalysts. However, the behaviour at the low potential region is inaccessible due to the redox peak at 1.38 ~ 1.41 V vs RHE overlapping with OER response (Figure 4(d)) and moreover the oxygen produced is difficult to determine [34,35]. Therefore, the real-time monitoring electrochemical mass spectrometry (EC-MS) was employed to investigate its electrocatalytic activity within a low potential region. The mass-specific current density was calculated by normalizing the detected oxygen mass current ($m/z = 32$) corresponding to OER based on the mass loading of the electrocatalyst. Figure 4(d) shows that

oxygen started to evolve at ~1.39 V vs RHE, reaching a specific current of 6.5 A/g at 1.44 V vs RHE. Instead for IrO₂, the onset potential for oxygen evolution was observed at a much higher potential of 1.48 V vs RHE. This indicates that, under the present synthesis and testing conditions, EL-4 displays more favorable OER kinetics than IrO₂. Based on the OER oxygen current derived from EC-MS, the Tafel slope of EL-4 at small current densities (<1 mA/cm²) was calculated to be 38.8 mV/dec, while this value is 43.6 mV/dec for IrO₂ (Figure 4(c)). Therefore, the exchange current density of EL-4 is 5.18×10^{-6} mA/cm², which is 20 times higher than that of IrO₂ (2.66×10^{-7} mA/cm²). This result again indicates the promising electrochemical properties of EL-4.

To investigate the structural origin of the high OER activity of EL-4, the microscopic morphology, element

composition, and surface valence states of the as-prepared EL-4 were characterized by scanning transmission electron microscopy (STEM) and energy dispersive X-ray spectroscopy (EDS) (Figure 5(a)). The cross-sectional high-angle annular dark field STEM (HAADF-STEM) image of the as-prepared EL-4 displays that the electrocatalyst film has an average thickness of 190 nm. The corresponding EDS element mapping images reveal the homogeneous distribution of Ru, Rh, V, Ni, W and O in the sample. Further investigation at high magnification (Figure 5(a), bottom panel) demonstrates the existence of small spherical particles in diameters less than 4 nm, which are composed of Rh, Ni and O elements. The high-resolution TEM (HRTEM) image in the inset of the HAADF-STEM image shows lattice fringes of 2.2 Å. Furthermore, the X-ray diffraction (XRD) pattern of EL-4 (Figure S3) reveals a reflection at the (200) plane, matching a NiO structure (JCPDS No. 47-1049). Therefore, the as-prepared EL-4 consists of rhodium-incorporated nickel oxide embedded in the multielement amorphous metal oxide. The XPS spectra (Figure S4) further confirm the presence of Ru, Rh, V, Ni, W and O in the material, which is essentially consistent with the HAADF-STEM elemental mapping results. The Rh element exists as Rh^{3+} , while the other metal elements (Ru, V, Ni, W) have the mixed valence states. In the O 1s XPS spectrum, the peaks at 530.4 and 531.8 eV correspond to lattice oxygen

(metal-bonded oxygen, M-O) and hydroxide groups (–OH), respectively.

Next, the morphology and compositions of EL-4 after electrochemical activation were investigated (Figure 5(b)). As multielement electrocatalysts transform their compositions and morphologies at the early stage of electrochemical reactions and evolve into active phases, we examined this activated phase of EL-4 in order to characterize the catalytically relevant state of the material. After the electrochemical activation step, the film thickness decreased to ~160 nm, and the corresponding elemental mapping indicates the dissolution of V during the activation. As depicted in the XPS survey spectrum (Figure S5), the V 2p peak disappears after activation, indicating the absence of this element on the surface. Moreover, Ru and W are no longer uniformly distributed in the sample but rather enriched their concentrations at the interface of the sample and the substrate (Figure 5(b), top panel, indicated as white arrows). The redistribution or dissolution of elements such as V and W can introduce oxygen vacancies, which not only create faster pathways for ion diffusion but also promote favorable OH-adsorption on the surface [36–38]. Additionally, the compositional rearrangement may contribute to electrochemical reconstruction, potentially leading to transformation or exposure of more active sites. The activated EL-4 retains the NiO phase (Figure S6). As shown in the high-magnification TEM images and

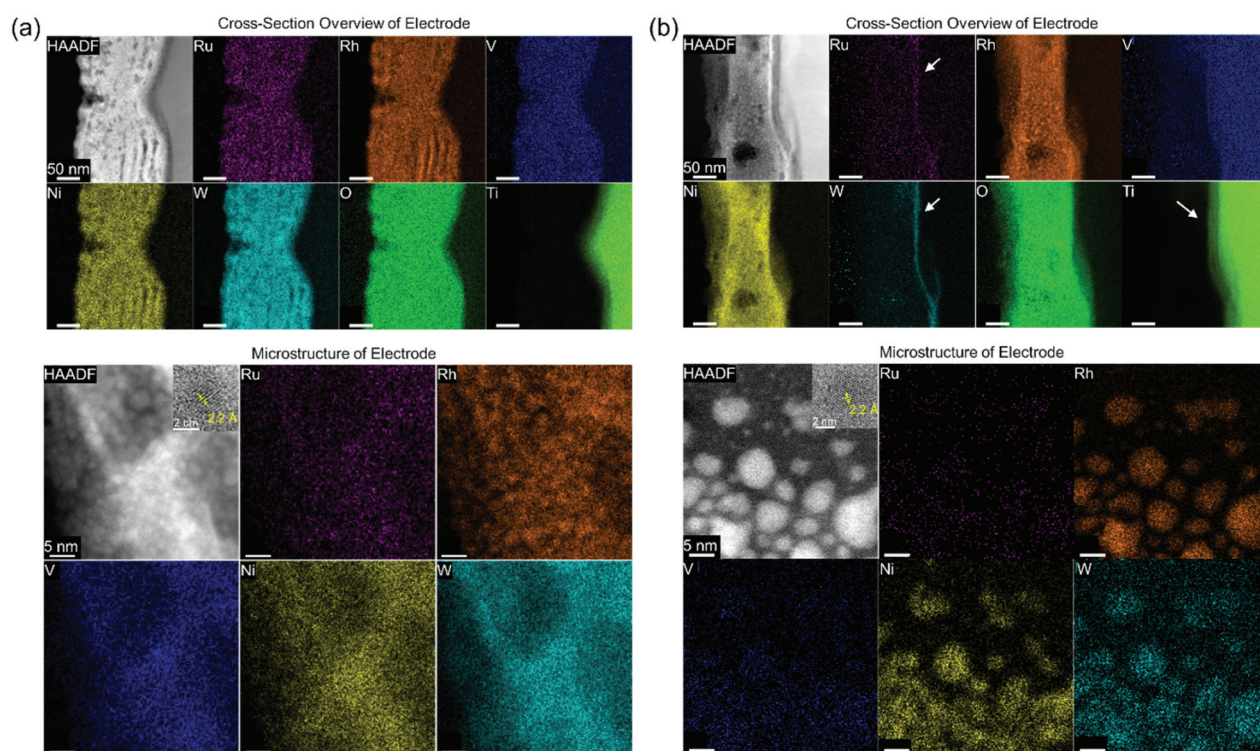


Figure 5. TEM observation of as-prepared (a) and activated EL-4 electrode (b). Top: cross-sectional HAADF-STEM images for overview of the electrode and the corresponding elemental mapping images. The scale bar is 50 nm. Bottom: HAADF-STEM microstructure and the corresponding elemental mapping images. The scale bar indicates 5 nm. Inset: HRTEM image of a representative single particle.

EDS mapping results in the bottom panel group of Figure 5(b), Rh, Ni, W and V are incorporated into larger nanoparticles, which suggests the diffusion of W and V into the rhodium nickel oxide structure during OER. The alloying of Ni, W, V and Rh modulates the local electronic structure of EL-4, thereby optimizing the adsorption energies of OER intermediates and improving electrocatalytic activities [39–42]. Figure S7 shows the high-resolution XPS spectra. The O 1s XPS spectrum is also deconvoluted into two peaks belonging to lattice oxygen and hydroxide. However, the contribution of hydroxide after the OER test greatly increased from 14% to 84%. The increase of the hydroxyl group is associated with the formation of oxygen vacancies during OER, which plays an important role in enhancing OER performance [37,43]. The transformation between the pristine material and the active phase poses a challenge for data-driven catalyst design. Traditional ML models predominantly rely on static descriptors of initial materials. However, our experimental results demonstrate that electrocatalytic performance depends on active-phase evolution under operating conditions. Therefore, integrating parameters that capture electrochemical reconstruction as dynamic ML descriptors is vital for accurately predicting true electrocatalytic performance.

To reveal the effect of metal elements on OER activity, we further prepared corresponding

quaternary-metal catalysts EL-4w/o M, where M denotes the absent metal element from EL-4 and checked their electrochemical properties. As shown in Figure 6(a), all the quaternary electrocatalysts show the worse OER performance compared to the quinary EL-4. When Rh or Ru is unavailable, the OER activity is significantly decreased. The overpotential values at 10 mA/cm² for EL-4w/oRh (= Ru_{0.5}V_{0.5}Ni_{0.5}W) and EL-4w/oRu (= Rh_{0.5}V_{0.5}Ni_{0.5}W) are 354.0 and 351.5 mV, respectively, which result shows an increase of more than 36 mV in comparison to EL-4 (Figure 6(b)). We put an emphasis here that EL-4w/oRh shows much lower current densities at high overpotentials than the other quaternary catalysts. Clearly, the OER activities of the quaternary compounds are inversely correlated with those of the unary ones (Figure S8). Their current densities at 1.54 V vs RHE are displayed in Figure 6(c). There is an opposite trend on current densities for unary oxides and their corresponding remaining-4-element oxides. When both Rh and Ru were removed, a ternary catalyst EL-4w/oRuRh (= V_{0.5}Ni_{0.5}W) was obtained. The electrocatalytic activity of EL-4w/oRuRh is further degraded, leading to a rather high overpotential of 520.8 mV at 10 mA/cm². The chronopotentiometric results (Figure 6(d)) also show that EL-4w/oRuRh is not as stable as the other quaternary and quinary materials. Based on the above results, the quinary electrocatalysts achieved enhanced OER activity and durability compared to quaternary

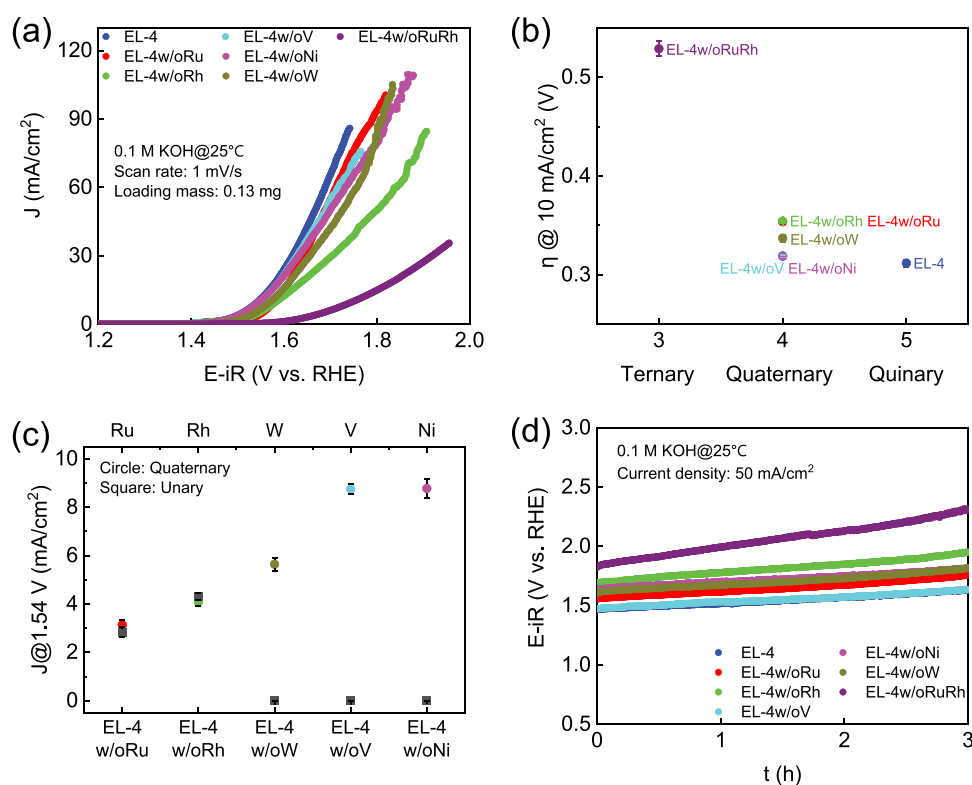


Figure 6. OER performances of multielement electrocatalysts measured in 0.1 M KOH electrolyte. (a) LSV curves measured at a scan rate of 1 mV/s. (b) Overpotentials at a current density of 10 mA/cm². (c) Current densities at 1.54 V vs RHE. (d) CP curves at a current density of 50 mA/cm² for 3 h.

and ternary materials. The excellent electrocatalytic performance of the quinary material should be attributed to the synergistic effect between multiple elements [44–46].

Conclusions

In conclusion, an interpretable data-driven framework using a small hybrid dataset has been developed for the extrapolative discovery of high-performance multielement oxygen evolution electrocatalysts from a large material search space, based on a small hybrid database composed of literature and experimental data. Our results identified a promising quinary oxide with a metal composition of Ru:Rh:V:Ni:W = 0.5:0.5:0.5:0.5:1. Electrochemical experiments demonstrated that this quinary oxide achieved a lower overpotential at 10 mA/cm² and showed improved alkaline oxygen evolution activity relative to IrO₂ and other oxides synthesized under the same condition. The consistency of experiments and predictions confirmed the effectiveness of our small-data-driven approach. Through further experiments, the quinary catalyst has been proved to exhibit higher OER activity than the oxides containing combinations of less metal elements due to multi-element synergistic effect. This highlights the necessity of the interpretable data-driven approach for the accelerated discovery of multielement metal oxide electrocatalysts and their understanding for better electrochemical properties.

Experimental

Feature calculation

To represent the composition of the electrocatalysts, 17 group coefficients and 17 group weighted properties (atomic radius in this case) were calculated as follows.

The group coefficient is derived from the total magnitude of the coefficient of a given element in the chemical formula of the electrocatalyst, if more than one element in the formula belongs to the same group, they are added. For example, in the formula Ba_{0.5}Co_{0.5}Fe_{1.5}O₅PrSr_{0.5} group coefficient 2 would equal 1 (Ba + Sr), while group coefficients 3, 8 and 9 would be 1 (Pr), 1.5 (Fe) and 0.5 (Co), respectively. For this study, all electrocatalysts in the curated data were deemed to be oxides, so oxygen was not counted in the calculation of group coefficient 16; thus it was deemed 0.

The weighted atomic radius was input as equal to the value of the atomic radius of the element in picometers. If more than one element belongs to the same group (such as Ba and Sr in the above formula) its value is calculated as a weighted average of them. Consequently, the weighted atomic radius for Ba_{0.5}

Co_{0.5}Fe_{1.5}O₅PrSr_{0.5} would be represented as 236 for group 2 (Ba + Sr, 0.5 × 253 + 0.5 × 219), and for groups 3, 8 and 9 it would be represented as 247 (Pr), 156 (Fe) and 152 (Co), if the electrocatalyst contains only oxygen for group 16, the atomic radius value was deemed 0.

Chemicals and materials

Nickel chloride (NiCl₂·6 H₂O, 98.0%), ruthenium chloride (RuCl₃, 99.9%), oxalic acid (H₂C₂O₄, 99%) and ethanol (99.5%) were purchased from FUJIFILM Wako, Japan. Tungsten chloride (WCl₆, ≥99.9%), iridium chloride (IrCl₃, 99.9%), rhodium chloride (RhCl₃, 99.95%), osmium chloride (OsCl₃, 99.9%), niobium chloride (NbCl₅, 99.995%) and potassium hydroxide (KOH, 99.99%) were obtained from Sigma-Aldrich, U.S.A. Vanadium chloride (VCl₃, min 95%) was purchased from Strem Chemicals, Inc., U.S.A., palladium chloride (PdCl₂, 99.999%) from Alfa Aesar, U.S.A., and acetone (99.5%) from KANTO KAGAKU, Japan. Milli-Q water (Merck KGaA, Germany; 18.2 MΩ·cm, TOC < 5 ppb) was used for preparing ultrapure electrolyte solutions. Titanium foil (Ti, 99.5%) was purchased from Nilaco, Japan. The foil was cut into 0.05 mm × 5 mm × 30 mm, cleaned by acetone, and then etched in H₂C₂O₄ aqueous solution (10%) at 80°C for 1 h. After washing with deionized water and dried, the etched Ti was used as substrate for electrocatalyst deposition.

Electrocatalyst synthesis

Electrocatalysts were synthesized using a dip-coating method described as the following. First, 0.1 M metal precursor solutions were prepared by dissolving metal salts in ethanol. A 0.05 M mixture was obtained by mixing the respective metal precursor solutions at a desired ratio, e.g. Ru : Rh : V : Ni : W = 0.5 : 0.5 : 0.5 : 0.5 : 1 for EL-4 (Ru_{0.5}Rh_{0.5}V_{0.5}Ni_{0.5}W). Then the mixed solution (2.5 μL) was dropped on an etched Ti substrate. A thin film was formed after drying at room temperature followed by heating at 350°C for 10 min. The deposition and heating process was performed successively for 10 times. Finally, the film was calcined at 450°C for 1 h, resulting in an oxide layer on the Ti substrate. For performance comparison, an IrO₂ film was also prepared using the same procedure.

Electrochemical measurements

The electrochemical measurements were performed on a multichannel potentiogalvanostat (VMP-3, Bio-Logic, France) in 0.1 M KOH aqueous solution at room temperature. In a standard three-electrode cell configuration, the as-prepared thin-film electrocatalysts were directly employed as the working electrodes.

A platinum wire and a Ag/AgCl electrode were used as the counter electrode and the reference electrode, respectively. The polarization curves were obtained by linear sweep voltammetry (LSV) at a scan rate of 1 mV/s, and the durability test was carried out by chronopotentiometry under a current density of 50 mA/cm². Cyclic voltammetry (CV) was measured between 0.6–0.8 V vs RHE at different scan rates of 5, 10, 20, 30 and 40 mV/s.

Electrochemical-mass spectroscopy measurements

Oxygen evolution was monitored on an electrochemistry mass spectrometry (EC-MS) system (Spectro Inlets ApS, Denmark). In this system, an electrochemical (EC) cell is coupled with a mass spectrometer (MS) by a membrane chip. The EC cell was installed, with a platinum disk coated with electrocatalysts as the working electrode, a platinum wire as the counter electrode and an Ag/AgCl electrode as the reference electrode, respectively. Measurements were carried out in 0.1 M KOH electrolyte, with a continuous flow of argon as carrier gas. EC data were obtained on VMP-3 and MS oxygen signals were collected using Zilien software.

Material characterization

X-ray diffraction (XRD) patterns were recorded on an X-ray diffractometer (Rigaku, Miniflex 600, Japan) with Cu K α radiation. Cross-sectional morphologies were observed using transmission electron microscopy/ scanning transmission electron microscopy (TFS Spectra Ultra TEM/STEM, Thermo Fisher Scientific, USA) operated at 300 kV, equipped with an energy-dispersive X-ray spectroscopy (EDS) detector for element analysis. Before observation, the sample was cut by focused ion beam (FIB-4000). X-ray photoelectron spectrometry (XPS) was carried out using an X-ray photoelectron spectrometer (PHI Quantes, ULVAC-PHI, Japan) with a monochromatic Al K α X-ray source (1486.6 eV). The obtained XPS spectra were calibrated using the C1s peak located at 284.8 eV.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the GteX Program Japan Grant Number [JPMJGX23H2], the MEXT Program: Data Creation and Utilization-Type Material Research and Development Project Grant Number [JPMXP1122712807], and JST COI-NEXT 'Center for Advanced Battery

Collaboration' Grant Number [JPMJPF2016]. This work was partially supported by Japan Science and Technology Agency (JST) as part of Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE), Grant Number [JPMJAP2421]. KS was indebted to Churchill College, Cambridge, and Cavendish Laboratory, the University of Cambridge with Visiting By-Fellowship, and National Institute for Materials Science.

References

- [1] Beyer L. Alwin Mittasch: industriechemiker und Philosoph. *Nachr Chem.* 2004 June 1;52(6):675–679. doi: [10.1002/nadc.20040520609](https://doi.org/10.1002/nadc.20040520609)
- [2] Mittasch A, Frankenburg W. Early studies of multi-component catalysts. In: Frankenburg WG, Komarewsky VI Rideal EK, editors. *Advances in catalysis*. Vol. 2. New York: Academic Press; 1950. p. 81–104.
- [3] Butler KT, Davies DW, Cartwright H, et al. Machine learning for molecular and materials science. *Nature.* 2018 July 1;559(7715):547–555. doi: [10.1038/s41586-018-0337-2](https://doi.org/10.1038/s41586-018-0337-2)
- [4] Sakaushi K, Watanabe A, Kumeda T, et al. Fast-decoding algorithm for electrode processes at electrified interfaces by mean-field kinetic model and Bayesian data assimilation: an active-data-mining approach for the efficient search and discovery of electrocatalysts. *ACS Appl Mater Inter.* 2022 May 25;14(20):22889–22902. doi: [10.1021/acsami.1c21038](https://doi.org/10.1021/acsami.1c21038)
- [5] Tom G, Schmid SP, Baird SG, et al. Self-driving laboratories for chemistry and materials Science. *Chem Rev.* 2024 Aug 28;124(16):9633–9732. doi: [10.1021/acs.chemrev.4c00055](https://doi.org/10.1021/acs.chemrev.4c00055)
- [6] Wang M, Ishii A, Sakaushi K. Accelerated electrocatalyst degradation testing by accurate and robust forecasting of multidimensional kinetic model with Bayesian data assimilation. *ACS Energy Lett.* 2025 Jan 10;10(1):22–29. doi: [10.1021/acsenergylett.4c02868](https://doi.org/10.1021/acsenergylett.4c02868)
- [7] Hoisang W, Sakaushi K. Key criteria for next-generation dimensionally stable electrodes towards large-scale green hydrogen production by water electrolysis. *Curr Opin Electrochem.* 2022 Dec 1;36:101136. doi: [10.1016/j.coelec.2022.101136](https://doi.org/10.1016/j.coelec.2022.101136)
- [8] Stamenkovic VR, Strmcnik D, Lopes PP, et al. Energy and fuels from electrochemical interfaces. *Nat Mater.* 2017 Jan 1;16(1):57–69. doi: [10.1038/nmat4738](https://doi.org/10.1038/nmat4738)
- [9] Ue M, Sakaushi K, Uosaki K. Basic knowledge in battery research bridging the gap between academia and industry. *Mater Horiz.* 2020;7(8):1937–1954. doi: [10.1039/D0MH00067A](https://doi.org/10.1039/D0MH00067A)
- [10] Sakaushi K. Science of electrode processes in the 21st century: fundamental understanding of microscopic mechanisms towards advancing electrochemical technologies. *Bull Chem Soc Jpn.* 2021;94(10):2423–2434. doi: [10.1246/bcsj.20210272](https://doi.org/10.1246/bcsj.20210272)
- [11] Wang M, Sakaushi K. Bridging classical concepts and modern advances in electrode processes: a tutorial review. *Phys Chem Chem Phys.* 2025;27(48):25743–25766. doi: [10.1039/D5CP03949E](https://doi.org/10.1039/D5CP03949E)
- [12] Yeh J-W, Chen S-K, Lin S-J, et al. Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. *Adv Eng Mater.* 2004;6(5):299–303. doi: [10.1002/adem.200300567](https://doi.org/10.1002/adem.200300567)

- [13] Yao Y, Huang Z, Xie P, et al. Carbothermal shock synthesis of high-entropy-alloy nanoparticles. *Science*. 2018;359(6383):1489–1494. doi: [10.1126/science.aan5412](#)
- [14] Huang Y-M, Chang C-W, Lin J-T, et al. Toward the Sabatier principle-guided design of low-platinum-group-metal trimetallic nanocatalysts for efficient hydrogen evolution and oxidation reactions. *Adv Funct Mater*. 2026;n/a(n/a):e14858.
- [15] Kao IT, Kuo R-T, Lin S-C, et al. Exploring a bimetallic catalyst family for hydrogen oxidation with insights into superior activity and durability. *Nat Commun*. 2025 Nov 25;16(1):10504. doi: [10.1038/s41467-025-65503-7](#)
- [16] Lin C-Y, He ZY, Lin J-T, et al. Atomically mixed high-entropy-alloy nanoframes with 3D subnanometer-thick electrocatalytic surfaces. *Adv Funct Mater*. 2025;35(45):2505927. doi: [10.1002/adfm.202505927](#)
- [17] Blundell TL, Sibanda BL, Sternberg MJE, et al. Knowledge-based prediction of protein structures and the design of novel molecules. *Nature*. 1987 Mar 1;326(6111):347–352. doi: [10.1038/326347a0](#)
- [18] Blundell TL, Mizuguchi K. Structural genomics: an overview. *Prog Biophys Mol Biol*. 2000 June 1;73(5):289–295. doi: [10.1016/S0079-6107\(00\)00008-0](#)
- [19] Baker D, Sali A. Protein structure prediction and structural genomics. *Science*. 2001;294(5540):93–96. doi: [10.1126/science.1065659](#)
- [20] Blundell TL, Jhoti H, Abell C. High-throughput crystallography for lead discovery in drug design. *Nat Rev Drug Discov*. 2002 Jan 1;1(1):45–54. doi: [10.1038/nrd706](#)
- [21] Senior AW, Evans R, Jumper J, et al. Improved protein structure prediction using potentials from deep learning. *Nature*. 2020 Jan 1;577(7792):706–710. doi: [10.1038/s41586-019-1923-7](#)
- [22] Castro Garcia A, Sakaushi K. Pitfalls in artificial intelligence powered discovery due to electrocatalyst evaluation methodologies. *ACS Electrochem*. 2025 Sep 4;1(9):1871–1877. doi: [10.1021/acselectrochem.5c00153](#)
- [23] Sakaushi K, Hoisang W, Tamura R. Human-machine collaboration for accelerated discovery of promising oxygen evolution electrocatalysts with on-demand elements. *ACS Cent Sci*. 2023 Dec 27;9(12):2216–2224. doi: [10.1021/acscentsci.3c01009](#)
- [24] Oviedo F, Ferres JL, Buonassisi T, et al. Interpretable and explainable machine learning for materials science and chemistry. *Acc Mater Res*. 2022 June 24;3(6):597–607. doi: [10.1021/accountsmr.1c00244](#)
- [25] Hartono NTP, Thapa J, Tiihonen A, et al. How machine learning can help select capping layers to suppress perovskite degradation. *Nat Commun*. 2020 Aug 20;11(1):4172. doi: [10.1038/s41467-020-17945-4](#)
- [26] Kotobi A, Singh K, Höche D, et al. Integrating explainability into graph neural network models for the prediction of X-ray absorption spectra. *J Am Chem Soc*. 2023 Oct 18;145(41):22584–22598. doi: [10.1021/jacs.3c07513](#)
- [27] Wu Y-J, Xu Y, Fang L, et al. Periodic table-based compositional descriptors for accelerating electrochemical material discovery: Li-ion conductors and oxygen evolution electrocatalysts. *Sci Technol Adv Mater*. 2025 Dec 31;5(1):2513218. doi: [10.1080/27660400.2025.2513218](#)
- [28] Kita H. Periodic variation of exchange current density of hydrogen electrode reaction with atomic number and reaction mechanism. *J Electrochem Soc*. 1966 Nov 1;113(11):1095. doi: [10.1149/1.2423772](#)
- [29] Trasatti S. Work function, electronegativity, and electrochemical behaviour of metals: III. Electrolytic hydrogen evolution in acid solutions. *J Electroanal Chem Interfacial Electrochem*. 1972 Sep 1;39(1):163–184. doi: [10.1016/S0022-0728\(72\)80485-6](#)
- [30] Taniike T, Takahashi K. The value of negative results in data-driven catalysis research. *Nat Catal*. 2023 Feb 1;6(2):108–111. doi: [10.1038/s41929-023-00920-9](#)
- [31] Rozemberczki B, Watson L, Bayer P, et al. The shapley value in machine learning. *arXiv E-Prints*. 2022: arXiv: 2202.05594.
- [32] Rossener Regonia P, Pelicano CM. Understanding the performance of (Ni–Fe–Co–Ce)ox-based water oxidation catalysts via explainable artificial intelligence framework. *ChemElectrochem*. 2024 May 2;11(9):e202300647. doi: [10.1002/celec.202300647](#)
- [33] Pitchai C, Huang C-F, Lo T-Y, et al. Machine learning-guided design of high-entropy FeCoCrMnCu layered double hydroxides for efficient oxygen evolution in alkaline media. *ACS Catalysis*. 2026 Mar 6;16(5):4502–4515. doi: [10.1021/acscatal.5c07303](#)
- [34] Scott SB, Rao RR, Moon C, et al. The low overpotential regime of acidic water oxidation part I: the importance of O₂ detection. *Energypamp Environ Sci*. 2022;15(5):1977–1987. doi: [10.1039/D1EE03914H](#)
- [35] Zheng Y-R, Vernieres J, Wang Z, et al. Monitoring oxygen production on mass-selected iridium–tantalum oxide electrocatalysts. *Nat Energy*. 2022 Jan 1;7(1):55–64.
- [36] Risch M, Grimaud A, May KJ, et al. Structural changes of cobalt-based perovskites upon water oxidation investigated by EXAFS. *J Phys Chem C*. 2013 May 2;117(17):8628–8635. doi: [10.1021/jp3126768](#)
- [37] Subramaniam T, Krishnaveni BS, Devaraj S. Tuning the surface oxygen vacancies of α -MnO₂ to enhance the kinetics of ORR and OER. *J Mater Sci*. 2024 June 20;35(17):1184. doi: [10.1007/s10854-024-12971-0](#)
- [38] Fernández-Climent R, Giménez S, García-Tecedor M. The role of oxygen vacancies in water splitting photoanodes. *Sustain Energy Fuels*. 2020;4(12):5916–5926. doi: [10.1039/D0SE01305F](#)
- [39] Ma Z, Wu K, Sun B, et al. Band engineering of AgSb1–xBiO3 for photocatalytic water oxidation under visible light. *J Mater Chem A*. 2015;3(16):8466–8474. doi: [10.1039/C5TA01020A](#)
- [40] Zhang B, Zheng X, Voznyy O, et al. Homogeneously dispersed multimetal oxygen-evolving catalysts. *Science*. 2016 Apr 15;352(6283):333–337. doi: [10.1126/science.aaf1525](#)
- [41] Xu S, Carter EA. Theoretical insights into heterogeneous (photo)electrochemical CO₂ reduction. *Chem Rev*. 2019 June 12;119(11):6631–6669. doi: [10.1021/acs.chemrev.8b00481](#)
- [42] Lee G, Jun SE, Kim Y, et al. Multicomponent metal oxide- and metal hydroxide-based electrocatalysts for alkaline water splitting. *Materials*. 2023;16(8):3280. doi: [10.3390/ma16083280](#)
- [43] Fabbri E, Nachtegaal M, Binnering T, et al. Dynamic surface self-reconstruction is the key of highly active perovskite nano-electrocatalysts for water splitting. *Nat Mater*. 2017 Sep 1;16(9):925–931. doi: [10.1038/nmat4938](#)

- [44] Lin Y, Tian Z, Zhang L, et al. Chromium-ruthenium oxide solid solution electrocatalyst for highly efficient oxygen evolution reaction in acidic media. *Nat Commun.* 2019 Jan 11;10(1):162. doi: [10.1038/s41467-018-08144-3](https://doi.org/10.1038/s41467-018-08144-3)
- [45] Liu D, Qiao L, Peng S, et al. Recent advances in electrocatalysts for efficient nitrate reduction to ammonia. *Adv Funct Mater.* 2023 Oct 1;33(43):2303480. doi: [10.1002/adfm.202303480](https://doi.org/10.1002/adfm.202303480)
- [46] Chen S, Liu D, Zhou P, et al. Multi-metal electrocatalyst with crystalline/amorphous structure for enhanced alkaline water/seawater hydrogen evolution. *J Colloid Interface Sci.* 2023 Nov 15;650:807–815. doi: [10.1016/j.jcis.2023.07.048](https://doi.org/10.1016/j.jcis.2023.07.048)