Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Engineering

Deeper learning in electrocatalysis: realizing opportunities and addressing challenges

John A Keith[1], James R McKone[1], Joshua D Snyder[2] and Maureen H Tang[2]

Emerging techniques in deep learning have created exciting opportunities for next-generation electrochemical technologies. While deep learning has been revolutionizing many research fields, strategies for its implementation for electrocatalysis remain nascent. This Opinion calls on the electrocatalysis community to join together and introduce a paradigm shift by establishing standards for reporting and sharing data from electrocatalysis investigations. We speculate on a possible future where crowd-sourced and standardized data from experimental and computational researchers can be analyzed collectively to better understand fundamental electrochemistry, yielding unprecedented insights for the development of new electrocatalysts. We identify key barriers to realizing this opportunity and how they might be overcome.

Addresses
[1] Department of Chemical and Petroleum Engineering Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15261, United States
[2] Department of Chemical and Biological Engineering, Drexel University, Philadelphia, PA 19104, United States

Corresponding author: John A Keith (jakeith@pitt.edu)

Current Opinion in Chemical Engineering 2022, **36**:100824

This review comes from a themed issue on **Reaction Engineering and Catalysis**

Edited by **Lars Grabow** and **Ryan Hartman**

https://doi.org/10.1016/j.coche.2022.100824

Introduction

Artificial intelligence and machine learning (AIML) have made important contributions for the development and applications in the chemical sciences [1•]. AIML allows the extraction of useful and insightful physical relationships from large and convoluted data sets. There are growing numbers of examples of AIML-guided design of materials [2,3] and molecules [4,5] where insights have been gleaned from large datasets comprising experimental and computational findings.

The high complexity of electrochemical environments and the large number of researchers producing data make electrocatalysis a ripe field for AIML. Electrocatalytic methods are widely considered to be attractive for generating fuels and chemicals from renewable feedstocks and driving global decarbonization [6•]. Moreover, efforts to address long-standing challenges in electrocatalysis are often enhanced by composition–activity relationships obtained through synergistic experimental and computational work [7•]. There is ample evidence that AIML tools are in place to advance the theories and practices of electrocatalysis, but we also recognize that common practices for experimental and computational research in this area are often inadequate for productive and insightful AIML studies.

One problematic issue pertains to the standards for how macroscale catalyst properties in their respective experiments are reported. For example, it is customary for measurements of electrochemical current flow to be converted to current densities by normalizing to one of several physical parameters, including the superficial area of the electrode, the microscopic or electrochemically-active surface area of the catalytic material, the mass of the catalyst used, or the number of putative active sites [8–10•]. Each of these treatments can yield useful information, but experimental reports often use only one convention and do not include enough information to convert to another normalization scheme. Similarly, one can often find inconsistent treatments of transport limitations, reference electrode calibrations, and ohmic resistance across the literature. Many of these inconsistencies can be identified and reconciled by trained practitioners, but it is much more difficult to do so algorithmically, which makes the associated data less valuable for AIML applications.

A second issue is the difficulty of measuring and controlling the actual chemical composition of a catalyst material interacting with its local environment. For example, it is difficult to generate atomically precise catalyst materials with well-controlled compositions, and it is also challenging to measure the precise composition of a practical electrocatalyst under ambient reaction conditions — this amounts to the electrochemical version of the well-known 'materials gap' [11] in heterogeneous catalysis. Similar challenges extend to understanding the nature of the local environment around the catalyst, for example, the atomic scale of electrode substrates and electrolytes, as well as the ways in which the experimental configuration influences observed reaction rates, for example, via nonuniform current densities across electrodes [12,13]. Accordingly, studies that deploy well-characterized analytical apparatuses and intentionally modulate the local environment around an electrocatalyst would be especially useful for AIML-integrated research.

A third issue revolves around data quality and curation. Within electrocatalysis, even simple statistical analyses such as replicates and error bars are routinely overlooked. Meanwhile, there is growing emphasis on ensuring that scientific data are FAIR (Findable, Accessible, Interoperable, Reusable) [14] Mendes et al. recently reviewed the challenges of applying FAIR principles to catalysis research [15]. They found that catalysis and, more generally engineering, greatly lagged biological sciences in applying FAIR data practices. Electrocatalysis may be more amenable to data-sharing practices than catalysis as a broader field because the current, voltage, and time are recorded directly by a potentiostat and do not require unique calibrations as, for example, a gas chromatograph does. Recent systematic efforts to apply FAIR data practices across European battery research laboratories support the amenability of electrochemistry to best practices for data storage and sharing [16].

To instigate a new epoch of AIML-driven electrocatalysis efforts, we call on the community to establish a new set of consistent and standardized expectations for how to report analytical data from experimental and computational electrocatalysis studies. Such an endeavor would require more effort than retaining the status quo, but we see significant opportunities where adopting uniform standards would enable the use of AIML to enhance and accelerate electrocatalysis research.

Big data in electrocatalysis: current practices and immediate opportunities

In contemporary electrocatalysis research, researchers report steady-state performance. For example, the oxygen reduction reaction (ORR) catalysis researchers generally report the kinetic current density per cm^2 on Pt at 0.9 V versus the RHE. Water-splitting catalyst studies often document the overpotential at 10 mA/cm^2 geometric current density. Reports on the CO_2 reduction reaction describe the faradaic selectivity for CO, H_2, and other products as a function of potential, where each data point is averaged over a fixed time interval — typically 1 hour. The 'holy grail' of AIML would use this type of performance data to develop models that predict optimal material properties from an array of compositions and structures, thus alleviating the need to synthesize and test many compounds. However, training data-driven models require large amounts of high-quality data across samples and researchers, which is challenging to acquire considering the intrinsic limitations of material synthesis and the practical issues described above.

A more tractable application of AIML in the near term is to use machine learning tools to classify reaction networks, extract kinetic parameters, and optimize reactor conditions using large datasets collected from a single study or research group. This approach can be seen as a natural extension of the types of continuum-scale models that chemical engineers routinely construct for electrochemical systems to diagnose limiting behavior, optimize parameters, and specify materials design. Indeed, researchers have found recently that applying AIML tools to such models of batteries [17•] and electro-organic synthesis [18] can fit kinetic and transport parameters more quickly and accurately than traditional methods.

An even greater opportunity for AIML in electrocatalysis lies in extracting useful information from data that are uniquely straightforward to collect with electrochemical apparatus, such as transient responses to electrical perturbations. Potential and current provide real-time control and measurement, respectively, of reaction rates. The ease of modulating potential and/or current, compared with temperature and/or pressure, opens the door to rapid acquisition of dynamic reaction data for model training. Incorporating generator–collector techniques such as the rotating ring–disk electrode or interdigitated electrode arrays [19] allow researchers to acquire megabytes or gigabytes of transient reaction data with very little additional effort compared with routine steady-state measurements. Addition of ancillary techniques such as differential or online electrochemical mass spectrometry can further provide complementary real-time chemical analysis.

Some examples of this transient approach can be found in the literature. Electrochemical noise analysis has been used to good effect in corrosion science [20], and these techniques are increasingly being deployed to analyze electrochemical energy conversion devices [21]. Oscillating potentials are known to impact Faradaic selectivity, and electrochemical parameters such as pulse frequency and duty cycle can optimize selectivity and/or conversion [22,23]. Recent interest in the e-refinery has renewed interest in dynamic voltammetry for energy conversion [24–26], and has even shown that a data-driven approach can predict the optimal reaction conditions [27••]. Outside the electrocatalysis field, electrochemical sensor arrays ('electronic tongues') have improved in their ability to classify the absence or presence of analytes by incorporating transient responses into training data [28]. Very recently, researchers applied deep learning to transient spectroscopy data to classify between 103 different kinetic models [29••]. Applying similar approaches to electrocatalysis systems could identify transport and kinetic mechanisms that lead to more efficient material development.

From the *in silico* arena, it is now commonplace for computational chemistry practitioners to be engaged in providing accompanying details about atomic scale reaction mechanisms. The past literature has seen many instances of purported computational validations of reaction mechanisms, but detailed justifications for the approximations themselves are harder to find. At best, we speculate that such details are often considered uninteresting and/or too specialized for a general audience, and thus they are not reported even though they should be. Journals are providing explicit guidance by setting baseline standards for how researchers should ideally demonstrate insights from computational research [30]. Computational data can be highly sensitive to the choice of model or underlying theoretical assumptions, and there is growing interest in explicitly understanding uncertainties associated with different computational procedures themselves [31–34].

Recommendations for cross-laboratory aggregation

Turning now to the broader challenge of harvesting data from numerous electrocatalysis research reports for AIML studies, we see two main challenges:

1. It is difficult to aggregate and make direct comparisons between different data sets due to disparities and ambiguities in data acquisition.
2. The statistical validity of existing literature data is often uncertain due to the lack of acquisition standards.

In theory, the solution to this is straightforward: there should be full disclosure of *all experimental and computational details* in a searchable SI document that includes rigorous statistical analysis. This is challenging at best to fully implement, but we note that computational groups have been implementing protocols that prioritize lucidity and reproducibility of computational data [35–38]. Here, we suggest a potential path forward.

The most glaring omission in the current electrocatalysis literature is a statistical standard for the reporting of electrocatalyst performance data. The problem this poses for AIML researchers is not knowing whether reported data are valid — that is, whether the materials, conditions and experimental and/or computational parameters were controlled and documented well enough for others to reproduce the result. This issue is made more critical when the data are gathered by trainees that may not have sufficient experience to know how to minimize error when making a measurement or when to exclude data. Specific 'engineering controls' have been previously proposed, including requiring new students to replicate a published experiment, data acquisition, or synthesis from their own group or another group within the field [39•–41]. This is a valuable practice that should be widely adopted, but it is not enforceable without a governing body equipped to certify the competency of every trainee. A more practical solution may be to develop best practices with respect to statistical rigor that can be adopted as requirements for publication. This includes setting a standard for the number of replicates that constitutes a statistically relevant data set and formulating standards that defines the bounds of relevance, that is, how do we identify and eliminate outliers without eliminating valuable data points? These standards could then be incentivized by journal editors requiring them. Although this may require a significant shift in culture for the electrocatalysis research community, we emphasize that this type of statistical rigor is already common practice in, that is, the biomedical field [42].

An example of a specific requirement might be that a viable data set must include at least the minimum number of replicates to produce an error bound that is not more than 50% of the measured value. Even better would be statistical treatments that explicitly differentiate the catalyst under study from a reasonable control. For example, can a researcher provide 95% confidence that a new binary catalyst is different in activity than either of its constituent components alone? We note that if the difference is large, as few as three measurements may be sufficient, but if the difference is very small, many dozens of measurements may be needed.

Another critical aspect here is to only remove true outliers but retain any measurements that might point toward new insights. There remains a risk, particularly for data used to report performance metrics, that excuses could be made to remove any data point that would lower the average value, artificially inflating reported average performance. For this reason, specific procedures for identifying outliers should be adopted. Two possibilities are using a Z-test and identifying outliers as those that are outside of a 95% confidence interval or using an interquartile range with graded outlier fences to identify minor and major outliers.

In addition to statistical relevance, the accuracy of any conclusions drawn from data harvesting is predicated on the viability of the comparisons being made. If the variability of the experimental or computational protocols used to obtain the mined data is too high, then any conclusions drawn from aggregating that data could be called into question. This can be remedied in part by requiring the detailed disclosure of all aspects of the experimental/computational protocols in a searchable supporting information document. Ideally, this information would be collected as a matter of course as an experiment is performed, for example, via standardized entries in electronic lab notebooks or within digital data files (analogous to CIF and VAMAS file formats for X-ray and surface science measurements, respectively). This will help to ensure that the data were acquired with the best practices and allow for categorization during data mining to ensure accurate comparisons. This can only go so far, however. Even with full disclosure, if mistakes were made that introduce systematic error or bias, then inaccuracies can be propagated during data aggregation even after accounting for random error with statistical analysis. A shift in the culture from an emphasis on novelty to one on reproducibility can help here. In electrocatalysis, common mistakes that can hurt data reproducibility and introduce systematic error are myriad and include: incorrect calibration of reference electrodes; incorrect measurement of catalyst mass or active area; deposition of precious metals from the counter electrode onto the working electrode; inaccurate measurement of catalyst composition; variability in ambient temperatures that are documented as ‘room temperature;’ incomplete saturation of electrolyte with reactant gas or incomplete deaerating of electrolyte; incorrect iR compensation; and potential sweep rates that are too fast. For computational data, reproducibility can likewise be impacted by typographical errors or unnoticed bugs in code. The fact that researchers are rarely obliged to provide all atomic-scale coordinates and/or openly provide codes and scripts used in data analysis does not help this cause. Improving reproducibility and the accuracy of aggregated comparisons requires the adoption of a standardized set of ‘best practices’ that will begin to bridge these experimental disparities. We have provided examples herein that can be used to motivate further discussion across the research community (Box 1 and Box 2).

These best practices are to: (1) ensure that data are obtained properly, (2) report all relevant parameters governing the acquisition of that data so that valid comparisons can be made, and (3) validate accuracy through reproducibility. Implementation of this, however, requires an overhaul of how data are acquired and how results are reported. The onus of much of this will fall on research journals via editorial practices and peer review. Here again, there is a clear path forward in the recent implementation of specific reporting/experimental requirements for publication, such as the battery checklists adopted by Joule and ACS Energy Letters [43]. These checklists guide researchers to adopt a valuable measure of uniformity in their practices. Organization of these reporting requirements into an AI-friendly SI format will significantly lower the barrier to mining data for AIML studies. Here, we take a first attempt at establishing an equivalent checklist for electrocatalysis, which we also submit for external scrutiny and discussion (Box 3).

Box 1 Experimental best practices

1. Voltammetric assessment of catalytic electrode before and after activity measurements
2. Calibration of reference electrode after running each measurement to be reported externally (include calibration data in SI)
3. Proper selection of counter electrode material and surface area
4. Slow sweep rates for activity measurements (i.e., ideally with validation data demonstrating equivalence to steady-state behavior)
5. Secondary measurements of catalyst loading and composition (e.g., via thermogravimetry or elemental analysis)
6. Statistically significant numbers of replicates
7. Post-mortem analysis of catalytic electrodes (TEM, XPS, XRD)
8. Accurate measure of pH and ionic concentration of electrolyte
9. Sufficient saturation of electrolyte with reactant or purge gas (validated with measurements of diffusion-limited currents or background currents, respectively)
10. Establishing specific figures of merit for specific electrocatalytic reactions
11. Ensure proper iR compensation by reporting measured cell resistance
12. All performance metrics for new materials should be compared to an established -IJstandard material;-İ e.g., Umicore or TKK Pt/C for oxygen reduction
13. Use of a fully documented break-in protocol until a steady-state response is obtained

Box 2 Computational best practices

1. Input files containing *all salient keywords that defined calculation methods employed.*
2. Salient XYZ coordinate files provided in a standard format with corresponding calculated energies and/or properties.
3. Appropriate demonstration of convergence of k-points and kinetic energy cutoffs when appropriate.
4. Appropriate demonstration that atomic-scale equilibration was reached during molecular simulations.
5. Explicit inclusion of any non-standard forcefields, auxiliary basis sets, and/or pseudopotential files.
6. Computational data and numerical analyses explicitly provided using clearly written code (e.g. Jupyter notebooks)

Box 3 Electrocatalysis checklist

Experimental details provided:

☐ Echem cell material
☐ Echem cell design
☐ Reference electrode type
☐ Counter electrode type
☐ Electrolyte composition (ionic concentration, pH)
☐ Method of reference electrode calibration (with calibration data included in the SI)
☐ Pre-conditioning protocol
☐ Electrode geometry (e.g. RDE, flowcell, h-cell, etc)
☐ Catalyst loading
☐ Catalyst binder material and loading
☐ Bulk catalyst composition
☐ Electrolyte temperature
☐ Magnitude of compensated series resistance

Computational details provided:

☐ Brief narrative summary of computational work (including software version numbers and hardware configurations used).
☐ Any non-commercial codes and/or code modifications that were used for data production and/or analysis.
☐ *All XYZ coordinates and/or unit cell parameters for any static structures or states mentioned in the text.*
☐ A reasonably truncated trajectory file consisting of structures obtained from Monte Carlo or molecular dynamics simulations.
☐ Explicit details about the calculation method, spin states / magnetic moments, system charge, etc.
☐ Copies of input files and auxiliary files with keywords used for simulations.
☐ Evidence of self-consistency of computational models (e.g. explicit data showing convergence in k-point sampling and kinetic energy cutoffs and/or equilibration of molecular simulations).

In summary, the straightforward and high-resolution nature of electrochemical measurements may be uniquely amenable to big-data approaches. However, realizing the promise of AIML in electrocatalysis requires aggregating data across researchers and laboratories with commonly accepted experimental and computational best practices to ensure data fidelity. A concerted effort by leaders in the research community, and particularly by journal editors, will help to incentivize the adoption of best practices and statistical rigor. Funding agencies also play a natural role in incentive structures and can facilitate efforts towards reproducibility and FAIR data practices. For example, standard metrics for reproducibility, in addition to existing metrics for mass activity or stability, could be created. Agencies could also provide resources specifically devoted to developing data infrastructures such as those described in Castelli et al. [16]. Ultimately, implementing such changes can enable deep learning techniques to reach their fullest potential for electrocatalytic reaction engineering and material discovery.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Support from the National Science Foundation (NSF) is acknowledged (CHE-1856460/1855657, CBET-1944192, DMR-1904571, CBET-2015859). JAK acknowledges support from the R.K. Mellon foundation, and JRM acknowledges support from the Arnold and Mabel Beckman Foundation.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest.

1. • Keith JA, Vassilev-Galindo V, Cheng B, Chmiela S, Gastegger M, Müller K-R, Tkatchenko A: **Combining machine learning and computational chemistry for predictive insights into chemical systems**. *Chem Rev* 2021, **121**:9816-9872.

An extensive and self-contained review of computational chemistry and machine learning methods.

2. Moosavi SM, Jablonka KM, Smit B: **The role of machine learning in the understanding and design of materials**. *J Am Chem Soc* 2020, **142**:20273-20287.

3. Liu Y, Zhao T, Ju W, Shi S: **Materials discovery and design using machine learning**. *J Materiomics* 2017, **3**:159-177.

4. Christensen M, Yunker LPE, Adedeji F, Häse F, Roch LM, Gensch T, dosPassosGomes G, Zepel T, Sigman MS, Aspuru-Guzik A, Hein JE: **Data-science driven autonomous process optimization**. *Commun Chem* 2021, **4**:112.

5. Haghighatlari M, Hachmann J: **Advances of machine learning in molecular modeling and simulation**. *Curr Opin Chem Eng* 2019, **23**:51-57.

6. • Schiffer ZJ, Manthiram K: **Electrification and decarbonization of the chemical industry**. *Joule* 2017, **1**:10-14.

Compact and informative overview of opportunities to mitigate carbon emissions in chemical manufacturing with electrochemistry.

7. • Seh ZW, Kibsgaard J, Dickens CF, Chorkendorff I, Nørskov JK, Jaramillo TF: **Combining theory and experiment in electrocatalysis: Insights into materials design**. *Science* 2017, **355**:eaad4998.

A comprehensive and authoritative overview of the current state of electrocatalysis research; excellent resource for new entrants.

8. Kibsgaard J, Chorkendorff I: **Considerations for the scaling-up of water splitting catalysts**. *Nat Energy* 2019, **4**:430-433.

9. Anantharaj S, Ede SR, Karthick K, SamSankar S, Sangeetha K, Karthik PE, Kundu S: **Precision and correctness in the evaluation of electrocatalytic water splitting: revisiting activity parameters with a critical assessment**. *Energy Environ Sci* 2018, **11**:744-771.

10. • Voiry D, Chhowalla M, Gogotsi Y, Kotov NA, Li Y, Penner RM, Schaak RE, Weiss PS: **Best practices for reporting electrocatalytic performance of nanomaterials**. *ACS Nano* 2018, **12**:9635-9638.

An insightful critique of the wide variability in reporting of electrocatalyst performance parameters.

11. **Mind the gap**. *Nat Catal* 2018, **1**:807-808 http://www.nature.com/articles/s41929-018-0188-0.

12. Wagner C: **Theoretical analysis of the current density distribution in electrolytic cells**. *J Electrochem Soc* 1951, **98**:116.

13. Parrish WR, Newman J: **Current distributions on plane, parallel electrodes in channel flow**. *J Electrochem Soc* 1970, **117**:43.

14. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, ’t Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B: **The FAIR Guiding Principles for scientific data management and stewardship**. *Sci Data* 2016, **3**:160018, , https://doi.org/10.1038/sdata.2016.18 number: 1 Publisher: Nature Publishing Group ⟨http://www.nature.com/articles/sdata201618⟩.

15. •• Mendes PSF, Siradze S, Pirro L, Thybaut JW: **Open data in catalysis: from today's big picture to the future of small data**. *ChemCatChem* 2021, **13**:836-850.

An assessment of the state of FAIR data practices within the broader catalysis community and identification of some technical challenges specific to catalysis and FAIR data practices.

16. •• Castelli IE, Arismendi-Arrieta DJ, Bhowmik A, Cekic-Laskovic I, Clark S, Dominko R, Flores E, Flowers J, UlvskovFrederiksen K, Friis J, Grimaud A, Hansen KV, Hardwick LJ, Hermansson K, Königer L, Lauritzen H, LeCras F, Li H, Lyonnard S, Lorrmann H, Marzari N, Niedzicki L, Pizzi G, Rahmanian F, Stein H, Uhrin M, Wenzel W, Winter M, Wölke C, Vegge T: **Data management plans: the importance of data management in the BIG-MAP project**. *Batter Supercaps* 2021, **4**:1803-1812.

Battery researchers from a large, multinational collaboration share the result of building a data infrastructure with intentional implementation of FAIR data principles. Figure 4 illustrates a potential dataflow for classifying experimental details in machine-readable format. The supporting information includes over 100 pages of deep detail on types of data and workflows for storage and access.

17. • Brady NW, Gould CA, West AC: **Quantitative parameter estimation, model selection, and variable selection in battery science**. *J Electrochem Soc* 2020, **167** 013501.

Researchers provide up-to-date and informative examples of how big-data techniques can be used for efficient parameter estimation in electrochemical systems (here, batteries).

18. Francis-Xavier F, Kubannek F, Schenkendorf R: **Hybrid process models in electrochemical syntheses under deep uncertainty**. *Processes* 2021, **9**:704.

19. Lee SE, Tang MH: **Asymmetric interdigitated electrodes for amperometric detection of soluble products**. *J Electrochem Soc* 2021, **168**:057519.

20. Cottis RA: **Interpretation of electrochemical noise data**. *Corrosion* 2001, **57**:265-285.

21. Denisov ES, Evdokimov YK, Martemianov S, Thomas A, Adiutantov N: **Electrochemical noise as a diagnostic tool for PEMFC**. *Fuel Cells* 2017, **17**:225-237.

22. Alkire R, Gould R: **Analysis of multiple reaction sequences in flow-? Through porous electrodes**. *J Electrochem Soc* 1976, **123**:1842.

23. Fedkiw PS, Scott WD: **Selectivity changes in electrochemical reaction sequences by modulated potential control**. *J Electrochem Soc* 1984, **131**:1304.

24. Kawamata Y, Hayashi K, Carlson E, Shaji S, Waldmann D, Simmons BJ, Edwards JT, Zapf CW, Saito M, Baran PS: **Chemoselective electrosynthesis using rapid alternating polarity**. *J Am Chem Soc* 2021, **143**:16580-16588.

25. Román AM, Spivey TD, Medlin JW, Holewinski A: **Accelerating electro-oxidation turnover rates via potential-modulated stimulation of electrocatalytic activity**. *Ind Eng Chem Res* 2020, **59**:19999-20010.

26. Gopeesingh J, Ardagh MA, Shetty M, Burke ST, Dauenhauer PJ, Abdelrahman OA: **Resonance-promoted formic acid oxidation via dynamic electrocatalytic modulation**. *ACS Catal* 2020, **10**:9932-9942.

27. •• Blanco DE, Lee B, Modestino MA: **Optimizing organic electrosynthesis through controlled voltage dosing and artificial intelligence**. *Proc Natl Acad Sci USA* 2019, **116**:17683-17689.

An illustrative example of how AI can be used to accelerate investigations of transient potential perturbations on catalytic activity, with direct application to industrially relevant chemical synthesis.

28. Cortina M, Duran A, Alegret S, delValle M: **A sequential injection electronic tongue employing the transient response from potentiometric sensors for anion multidetermination**. *Anal Bioanal Chem* 2006, **385**:1186-1194.

29. •• Kollenz P, Herten D-P, Buckup T: **Unravelling the kinetic model of photochemical reactions via deep learning**. *J Phys Chem B* 2020, **124**:6358-6368.

Large two-dimensional datasets of wavenumber versus time were numerically simulated for 103 different kinetic mechanisms of photochemical reactions; these datasets were used to train a neural network model that then predicted the probability that experimental data obeyed each mechanism.

30. Schneider WF: **New physical insights from a computational catalysis perspective**. *J Phys Chem C* 2017, **121**:15491-15492.

31. Zhan N, Kitchin JR: **Uncertainty quantification in machine learning and nonlinear least squares regression models**. *AIChE J* 2021,e17516, , https://doi.org/10.1002/aic.17516

32. Wang A, Kingsbury R, McDermott M, Horton M, Jain A, Ong SP, Dwaraknath S, Persson KA: **A framework for quantifying uncertainty in DFT energy corrections**. *Sci Rep* 2021, **11**:15496.

33. Gabriel JJ, Paulson NH, Duong TC, Tavazza F, Becker CA, Chaudhuri S, Stan M: **Uncertainty quantification in atomistic modeling of metals and its effect on mesoscale and continuum modeling: a review**. *JOM* 2021, **73**:149-163.

34. Maldonado AM, Hagiwara S, Choi TH, Eckert F, Schwarz K, Sundararaman R, Otani M, Keith JA: **Quantifying uncertainties in solvation procedures for modeling aqueous phase reaction mechanisms**. *J Phys Chem A* 2021, **125**:154-164.

35. Huber SP, Zoupanos S, Uhrin M, Talirz L, Kahle L, Häuselmann R, Gresch D, Müller T, Yakutovich AV, Andersen CW, Ramirez FF, Adorf CS, Gargiulo F, Kumbhar S, Passaro E, Johnston C, Merkys A, Cepellotti A, Mounet N, Marzari N, Kozinsky B, Pizzi G: **AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance**. *Sci Data* 2020, **7**:300.

36. Álvarez-Moreno M, De Graaf C, López N, Maseras F, Poblet JM, Bo C: **Managing the computational chemistry big data problem: the ioChem-BD platform**. *J Chem Inf Model* 2015, **55**:95-103.

37. Govoni M, Munakami M, Tanikanti A, Skone JH, Runesha HB, Giberti F, de Pablo J, Galli G: **Qresp, a tool for curating, discovering and exploring reproducible scientific papers**. *Sci Data* 2019, **6**:190002.

38. Kitchin JR: **Examples of effective data sharing in scientific publishing**. *ACS Catal* 2015, **5**:3894-3899.

39. • Bender G, Carmo M, Smolinka T, Gago A, Danilovic N, Mueller M, Ganci F, Fallisch A, Lettenmeier P, Friedrich K, Ayers K, Pivovar B, Mergel J, Stolten D: **Initial approaches in benchmarking and round robin testing for proton exchange membrane water electrolyzers**. *Int J Hydrog Energy* 2019, **44**:9174-9187.

Study identifies level of uncertainty that arises simply from testing of a given membrane electrode assembly by different lab groups.

40. Ritter S, Bosch R-W, Huet F, Ngo K, Cottis RA, Bakalli M, Curioni M, Herbst M, Heyn A, Macak J, Novotny R, Öijerholm J, Saario T, Sanchez-Amaya JM, Takenouti H, Zajec B, Zhang W: **Results of an international round-robin exercise on electrochemical impedance spectroscopy**. *Corros Eng Sci Technol* 2021, **56**:254-268.

41. Siegmund D, Metz S, Peinecke V, Warner TE, Cremers C, Grevé A, Smolinka T, Segets D, Apfel U-P: **Crossing the valley of death: from fundamental to applied research in electrolysis**. *JACS Au* 2021, **1**:527-535.

42. Hardwicke TE, Goodman SN: **How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey**. *PLoS One* 2020, **15**:e0239598.

43. Sun Y-K: **An experimental checklist for reporting battery performances**. *ACS Energy Lett* 2021, **6**:2187-2189.