



U.S. Department of Energy
Advanced Research Projects Agency - Energy (ARPA-E)

Request for Information
DE-FOA-0003375

on

Accelerating the Catalyst Development Cycle for Net Zero Applications

Introduction:

The purpose of this Request for Information (RFI) is to solicit input for a potential ARPA-E program focused on accelerating the heterogeneous catalyst development cycle for incorporation into reactors, devices, equipment, unit operations, and process technology applications relevant to the U.S. 2050 net zero goals. These material development cycles can take decades to complete, starting from the discovery scale at milligram quantities and finishing at the development scale with kilogram quantities. ARPA-E is interested in decreasing the length of development cycles (Figure 1) to months while capturing significant energy efficiency increases, emissions reductions, and/or precious metal reductions. Major bottlenecks in the process include inefficient discovery, irreproducible multi-scale synthesis, laborious characterization, narrow design space optimization, irrelevant performance evaluation, and impractical integration (i.e., not “drop-in”) of heterogeneous catalysts into emerging technologies.

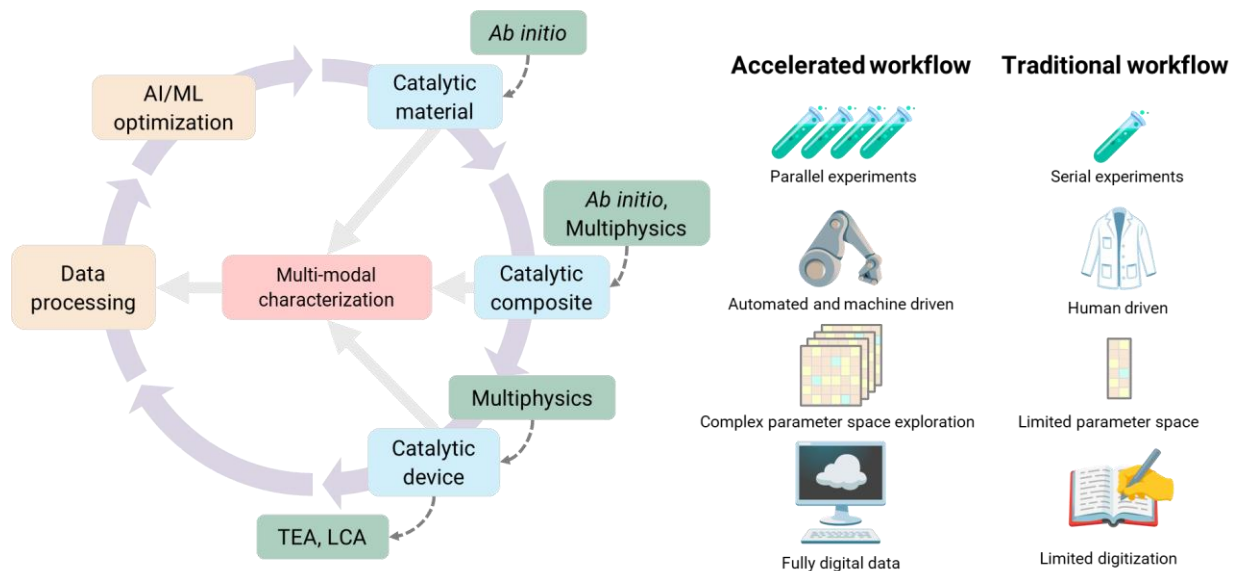


Figure 1. Envisioned accelerated development cycle and comparison with a traditional workflow, including artificial intelligence (AI), machine learning (ML), techno-economic analysis (TEA), and life cycle analysis (LCA).



ARPA-E is interested in identifying potentially disruptive techniques or workflows that expedite:

- Integration of catalytic material discovery and synthesis with device (i.e., cell or reactor) performance in a rapid, parallel, automated, and/or combinatorial manner. Devices should operate under realistic working conditions and correlate to deployment at relevant scales.
- Utilization of hardware automation and modern data science to generate, handle, and process large quantities of high quality, multi-dimensional experimental data.

Such approaches must ultimately accelerate the optimization and feedback at each level of complexity from material synthesis to device.

To facilitate program development, ARPA-E is seeking input from experts as described in Table 1 below.

Area	Subject Matter Expert Examples
Catalysis	<ul style="list-style-type: none">• Materials scientists• Chemical reaction engineers• Electrochemists• Thermochemists• Characterization experts• Process engineers• Catalysis experts (from atomic scale to commercial manufacturers)
Acceleration	<ul style="list-style-type: none">• Automation experts• Robotics experts• Self-driving laboratory experts• High throughput experimentalists• Parallel testing experts
Computation	<ul style="list-style-type: none">• Modeling experts (density-functional theory, molecular dynamics, Monte Carlo, multiphysics, etc.)• Data scientists/statisticians• AI, ML, and computer science experts• Data management experts

The tools and workflows of interest should be generalizable and applied to catalyst and process optimization activities across the same class of catalytic chemistry (e.g., electrochemical or thermochemical) that significantly impacts energy technologies of interest to ARPA-E. The questions in this document are to assist in providing input on the following topics:

- **Advancements and limitations in hardware:** Automation, combinatorial experimentation, and parallelization can establish an accelerated workflow that covers material discovery, synthesis, characterization, processing, and integration into functional devices/reactors/units. ARPA-E is interested in identifying major rate limiting and bottlenecking steps in traditional catalyst research and development (R&D) workflows that may benefit from new hardware.
- **Advancements and limitations in software:** Computer-based tools can automate and accelerate data processing, hypothesis generation, experimental design, and/or conclusion formation. ARPA-E is interested in statistical, machine learning, and/or artificial intelligence algorithms and tools, data requirements, database generation and management, closed-loop and feed-forward



advanced process control, computational power, (in)compatibility with existing open-source data, and future software advancement compatibility with catalyst R&D workflows.

- **Tightly coupled experimentation and theory:** ARPA-E is interested in novel workflow approaches which can integrate theoretical predictions, simulations, or computational modeling with experimentally determined empirical results. This includes the efficacy of such approaches, along with resources required for implementation.
- **Success metrics:** ARPA-E is interested in how to quantify the impact of new automated tools and workflow approaches for catalytic material-to-device optimization, as well as potential limitations, risks, and risk management approaches.

Areas Not of Interest for Responses to this RFI:

- Work focused on basic research aimed purely at fundamental knowledge generation.
- Experimental catalysis outside of electrochemical and thermochemical systems, including:
 - Homogeneous catalysis;
 - Plasma catalysis;
 - Photocatalysis; and
 - Battery chemistry (electric vehicle applications).
- Work focused purely on generating synthetic data.

RFI Guidelines:

CAREFULLY REVIEW ALL RFI GUIDELINES BELOW.

Note that the information you provide will be used by ARPA-E solely for program planning, without attribution. **THIS IS A REQUEST FOR INFORMATION ONLY. THIS NOTICE DOES NOT CONSTITUTE A FUNDING OPPORTUNITY ANNOUNCEMENT (FOA). NO FOA EXISTS AT THIS TIME.**

The purpose of this RFI is solely to solicit input for ARPA-E consideration to inform the possible formulation of future research programs. ARPA-E will not provide funding or compensation for any information submitted in response to this RFI, and ARPA-E may use information submitted to this RFI without any attribution to the source. This RFI provides the broad research community with an opportunity to contribute views and opinions.

No material submitted for review will be returned and there will be no formal or informal debriefing concerning the review of any submitted material. ARPA-E may contact respondents to request clarification or seek additional information relevant to this RFI. All responses provided will be considered, but ARPA-E will not respond to individual submissions or publish a compendium of responses. **Respondents shall not include any information in the response to this RFI that could be considered proprietary or confidential.**

Responses to this RFI should be submitted in PDF format to the email address **ARPA-E-RFI@hq.doe.gov** by **5:00 PM Eastern Time on Thursday, June 13, 2024.**



Emails should conform to the following guidelines:

- Insert “<your organization name>” - Response to Accelerating the Catalyst Development Cycle for Net Zero Applications” in the email subject line.
- In the body of your email, include your name, title, organization, type of organization (e.g., university, non-governmental organization, small business, large business, federally funded research and development center [FFRDC], government-owned/government-operated [GOGO]), email address, telephone number, and area of expertise.
- In the body of your email, note which question(s) you are answering using the provided format (e.g., I.a, II.b).
- Responses to this RFI are limited to no more than 10 pages in length (12-point font size).
- Responders are strongly encouraged to include preliminary results, data, and figures that describe their potential materials, designs, or processes.

Technical Background:

2050 Net Zero Emissions Goals

To achieve the U.S. 2050 net zero goals, the DOE has set major energy and emissions goals for establishing a hydrogen and carbon economy. This includes generating 50 megatonnes (Mt) of clean hydrogen, with emissions less than 0.45 kilogram (kg) of carbon dioxide (CO₂) equivalent per kg of hydrogen (H₂) (kgCO₂e/kgH₂), and sequestering gigatonnes (Gt) of CO₂ per year by 2050.^{1,2} Currently, the U.S. produces less than 1 Mt/year of clean hydrogen and captures CO₂ at a rate lower than 1 Mt/year.^{1,3} Concurrently, global consumption of fossil resources, such as crude oil and natural gas, will either maintain similar levels (greater than 100 million barrels per day [MM bpd]) as today, or require significant reduction (less than 25 MM bpd) to achieve the net zero emissions scenario.^{4,5} As renewable electricity and electric vehicles become more abundant, petrochemicals can become the largest driver for crude oil demand.⁶ A transition to an economic hydrogen and carbon economy will critically involve electrochemical and thermochemical catalytic processes to produce the low emission chemicals, fuels, and materials by 2050. Presently, many proposed catalytic reactions are inefficient and require significant electricity, heat, and critical materials.

¹ Hannah Murdoch et al., “Pathways to Commercial Liftoff, Clean Hydrogen” (US Department of Energy, 2023).

² “Carbon Negative Shot,” Energy.gov, <https://www.energy.gov/fecm/carbon-negative-shot>.

³ “Regional Direct Air Capture Hubs,” Energy.gov, <https://www.energy.gov/oced/DACHubs>.

⁴ Joseph DeCarolis and A LaRose, “Annual Energy Outlook 2023” (US Energy Information Administration, 2023).

⁵ “Oil and Gas in Net Zero Transitions” (IEA, 2023), <https://www.iea.org/reports/the-oil-and-gas-industry-in-net-zero-transitions>.

⁶ “The Future of Petrochemicals” (IEA, 2018), <https://www.iea.org/reports/the-future-of-petrochemicals>.



Examples of major challenges in future catalytic reactions are given below:

- **Electrochemical challenges:** Electrochemistry is an attractive solution that could take advantage of a low-emission grid to produce fuels, commodity chemicals, and raw materials using only electricity.^{7,8} Presently hydrogen production, CO₂ reduction, and other proposed electrochemical reactions are based on variations of the water electrolysis reaction.^{8,9} However, the energy requirements for water electrolysis (approximately 50 kilowatt-hour per kg H₂ [kWh/kgH₂]) are substantially higher than the output energy of hydrogen (higher heating value of 39 kWh/kgH₂).^{10,11} In addition, the oxygen evolution reaction at the anode often uses iridium as the catalytic material, which is constrained by limited supply chains.¹² Thus, optimization for increased efficiency, decreased use of critical minerals, and improved costs are of significant interest.¹³
- **Thermochemical challenges:** Thermochemistry has an established infrastructure with catalysts well-optimized for current crude oil and natural gas-based feedstocks.¹⁴ However, the carbon transition means that there is an increased need for non-traditional carbon sources (e.g., biogenic, recycled, captured) for fuel feedstocks (Figure 2).^{14,15} These fuel outputs, which represent 28 quadrillion British thermal units of energy in the U.S., will require re-optimization of catalysts to utilize new forms of carbon.^{16,17} In addition, this also applies to the decarbonization of high-value petrochemicals, which are currently economic due to fuel production.⁶ Finally, this can also be applied to the CO₂ capture processes, where significant improvements are needed for sorbents to make CO₂ capture feasible and economic.¹⁸

⁷ Jordi Cabana et al., “NGenE 2021: Electrochemistry Is Everywhere,” *ACS Energy Letters* 7, no. 1 (January 14, 2022): 368–74.

⁸ Hasan Ozcan et al., “Recent Advances, Challenges, and Prospects of Electrochemical Water-Splitting Technologies for Net-Zero Transition,” *Cleaner Chemical Engineering* 8 (December 1, 2023): 100115.

⁹ Marcelo Carmo et al., “A Comprehensive Review on PEM Water Electrolysis,” *International Journal of Hydrogen Energy* 38, no. 12 (April 22, 2013): 4901–34.

¹⁰ “Technical Targets for Proton Exchange Membrane Electrolysis,” Energy.gov, <https://www.energy.gov/eere/fuelcells/technical-targets-proton-exchange-membrane-electrolysis>.

¹¹ Johanna Ivy Levene et al., “An Analysis of Hydrogen Production from Renewable Electricity Sources,” *Solar Energy* 81, no. 6 (June 2007): 773–80.

¹² Julia E. Greenwald, Mervin Zhao, and Douglas A. Wicks, “Critical Mineral Demands May Limit Scaling of Green Hydrogen Production,” *Frontiers in Geochemistry* 1 (January 16, 2024): 1328384.

¹³ Alex Badgett et al., “Updated Manufactured Cost Analysis for Proton Exchange Membrane Water Electrolyzers,” February 16, 2024.

¹⁴ Avinash Alagumalai, Balaji Devarajan, and Hua Song, “Unlocking the Potential of Catalysts in Thermochemical Energy Conversion Processes,” *Catalysis Science & Technology* 13, no. 19 (2023): 5632–53.

¹⁵ Yixiao Wang et al., “Catalytic Processes to Accelerate Decarbonization in a Net-Zero Carbon World,” *ChemSusChem* 15, no. 24 (December 20, 2022): e202201290.

¹⁶ “U.S. Energy Facts Explained - Consumption and Production - U.S. Energy Information Administration (EIA),” <https://www.eia.gov/energyexplained/us-energy-facts/>.

¹⁷ Muhammed Zafar Ali Khan et al., “Potential of Clean Liquid Fuels in Decarbonizing Transportation – An Overlooked Net- Zero Pathway?,” *Renewable and Sustainable Energy Reviews* 183 (September 2023): 113483.

¹⁸ Bartosz Dziejarski et al., “CO₂ Capture Materials: A Review of Current Trends and Future Challenges,” *Materials Today Sustainability* 24 (December 2023): 100483.

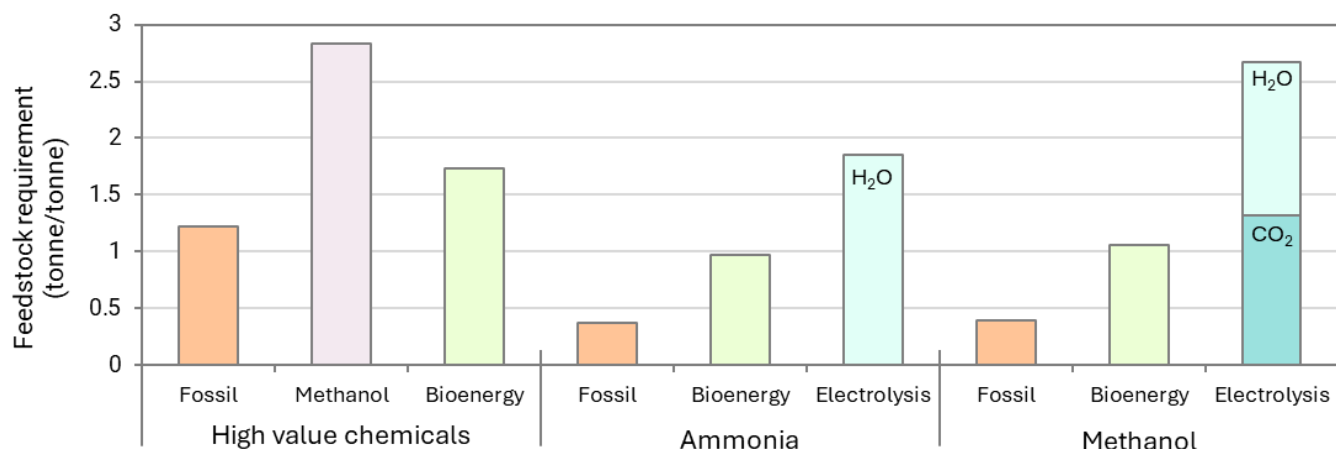


Figure 2. Relative input requirements for chemicals based on feedstock type. Transitioning from fossil feedstock to bioenergy and electrolysis will require higher inputs of new feedstock.¹⁹

Challenge of Bridging Traditional Multi-Scale Catalyst Development Workflows

Traditional catalysis R&D workflows at the lab-scale are based on rational surface science, synthesis, and characterization of new materials. Lab-scale R&D synthesis is typically in the sub-gram to gram range and tested over relatively short timescales of hours to days under various environmental and operational conditions. This can lead to hypothesis, modeling, and structure-property relationships. Outputs are typically scientific literature focused on theoretical discovery and experiments with variable composition or nanoscale morphology. Unfortunately, many of these prepared materials are too risky or expensive to commercialize, and the workflow takes too long for industry partners (e.g., technology licensors, catalyst original equipment manufacturers) to adopt. Often, this R&D cycle can take years, as it is focused on basic scientific discovery.^{20,21}

On the other hand, industrial catalysis researchers follow technology development workflows on the order of months to years and investigate device or unit reactor performance. Whereas catalysis R&D often tests the pure or “neat” catalytic materials, industrial catalysis focuses on the engineered-formed known catalytic composites with variable active catalyst loading. These composites range from catalyst-coated electrochemical membranes to thermochemical reactor extrudate, pellets, and monoliths. Critically, the additives (e.g., stabilizers, lubricants, ionomers) that bind the catalytic material in the composite also decrease the availability of active sites and reduce catalytic activity.^{22,23} Industrial catalysts are also synthesized and utilized in kilograms or tonnes, while operations need to be

¹⁹ Adapted from “The Future of Petrochemicals” (IEA, 2018), <https://www.iea.org/reports/the-future-of-petrochemicals>. Creative Commons License (CC BY 4.0), <https://creativecommons.org/licenses/by/4.0/>.

²⁰ Joshua A. Schaidle et al., “Transitioning Rationally Designed Catalytic Materials to Real ‘Working’ Catalysts Produced at Commercial Scale: Nanoparticle Materials,” in *Catalysis*, ed. James Spivey and Yi-Fan Han, vol. 29 (Cambridge: Royal Society of Chemistry, 2017), 213–81.

²¹ Thomas W Eagar, “Bringing New Materials to Market,” *Technology Review* 98, no. 2 (1995): 42–49.

²² Konrad Ehelebe et al., “Benchmarking Fuel Cell Electrocatalysts Using Gas Diffusion Electrodes: Inter-Lab Comparison and Best Practices,” *ACS Energy Letters* 7, no. 2 (February 11, 2022): 816–26.

²³ Gareth T. Whiting et al., “Multiscale Mechanistic Insights of Shaped Catalyst Body Formulations and Their Impact on Catalytic Properties,” *ACS Catalysis* 9, no. 6 (June 7, 2019): 4792–4803.

performed over months (greater than 2000 hours) and heat and mass transfer effects are more pronounced.²⁴

Thus, there exists a large gap in the length and testing timescales for which the two development cycles exist. This is exacerbated by the siloing of R&D between these two groups, as well as the distinct performance metrics between catalytic particles (focused on activity) and catalytic composites (focused on stability). In addition, the different length scales involve significantly different physics, as the geometry, configuration, and operating principles of the lab-scale test reactor versus the pilot-scale can be significantly distinct. Transitioning from catalytic material-to-device leads to a significant and complex parameter space evolution. For example, the lab-scale, active nanostructured catalysts can fall short in aspects of material durability, stability, synthesis feasibility, or operating resilience. Conversely, industrial scale researchers take incremental risks and may not explore wide parameter or configuration spaces, as this intrinsically increases risk and need for re-optimization.

The notion of re-inventing these separate conventional workflows at each scale (Figure 3) and combining them all into a single workflow may present an opportunity to accelerate the whole catalyst design and development cycle on the order of months instead of years.^{25,26}

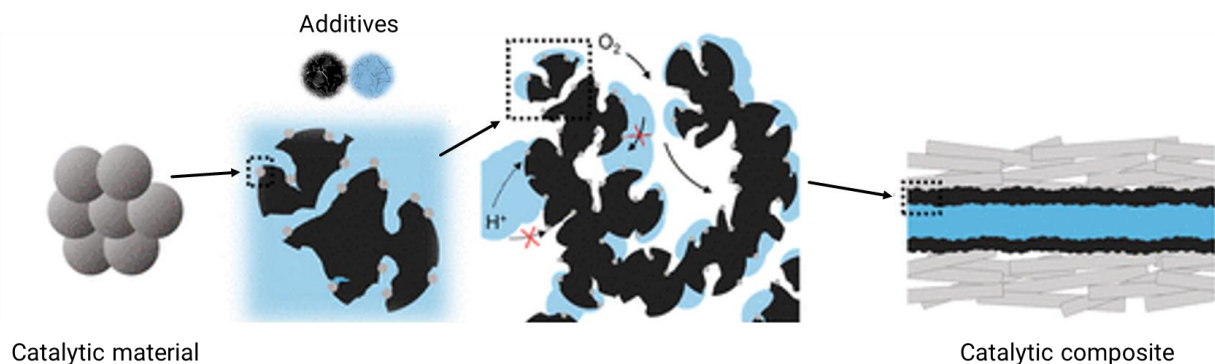


Figure 3. Transformation of active catalytic nanomaterials (e.g., nanoparticles/powders) into composite engineered forms.²⁷

Technical Approaches to Overcome Traditional Workflow Challenges

ARPA-E has identified two major technical thrusts that may accelerate the catalysis development workflow:

1. The first approach is to use integrated hardware and software innovations to create accelerated workflows with multiple degrees of freedom (i.e., not only elemental composition) that organize

²⁴ Susannah L. Scott, ed., “A Matter of Life(Time) and Death,” *ACS Catalysis* 8, no. 9 (September 7, 2018): 8597–99.

²⁵ Adrian Ramirez et al., “Accelerated Exploration of Heterogeneous CO₂ Hydrogenation Catalysts by Bayesian-Optimized High-Throughput and Automated Experimentation,” *Chem Catalysis* 4, no. 2 (February 15, 2024).

²⁶ Marçal Capdevila-Cortada, “Closed-Loop Optimization,” *Nature Catalysis* 7, no. 2 (February 27, 2024): 114–114.

²⁷ Adapted from Konrad Ehelebe et al., “Benchmarking Fuel Cell Electrocatalysts Using Gas Diffusion Electrodes: Inter-Lab Comparison and Best Practices,” *ACS Energy Letters* 7, no. 2 (February 11, 2022): 816–26. Adapted with permission.



into a single, multi-scale active catalytic material-to-device stream. For example, as opposed to a traditional combinatorial system-based workflow which could simultaneously synthesize 1,000 catalytic active material candidates at sub-microgram scale and test for their catalytic activity in proxy-reactors or rotating disk electrodes, the target system and workflow proposed here would be able to measure or correlate the performance of 1,000 monolith-level or membrane electrode assembly (MEA)-level catalytic devices. The challenge is both in the varying form factor of the active catalytic material (nanoparticle versus composite), as well as the distinct environment conditions that impact both heat and mass transfer.

Due to the multi-dimensional aspect of the parameter space, advances in high throughput experimentation, parallelization, statistical and ML tools, and robotics are expected to enable a feedback loop for optimizing these new catalytic devices. These lab-scale reactors/devices/proxy-devices must have a high-degree of correlation with commercial scale performance by at least capturing the most important variables for scale-up operation. It should be noted that the combinatorial composition screening is an acceptable task node and could be included in this multi-scale workflow as a guide for assisting with material discovery.

2. The second approach is to develop the ability to integrate theoretical predictions, simulations, and modeling into critical aspects of the accelerated multi-scale experimental workflow above. While there have been independent efforts to use data science, ML, and AI tools for optimizing processes based on simulations and theory and for analyzing experimental data, there has not been a widespread, concerted effort of creating a synergistic approach of studying experimental multi-dimensional parameter space guided by physics- or chemistry-based models. If successful, this effort can yield new insight into generalizing the workflow, as well as uncover new mechanisms which may fundamentally change the understanding of certain catalytic processes.

ARPA-E acknowledges that challenges exist in the lack of widespread availability and accessibility of the tools necessary for the hardware and analytical infrastructure governing a workflow, such as material characterization bottlenecks, incompatible device fabrication, or post-processing methods. In addition, it is difficult to reconcile open source and literature data with experimentally generated data, as the specific conditions cannot always be well-correlated. Further, literature results are typically presented with a positive bias excluding access to the entire raw data set.

RFI Questions:

The questions posed in this section are classified into several different groups as appropriate. Respondents may provide responses and information about any of the following questions. **ARPA-E does not expect any one respondent to answer all, or even many, of the prompts in this RFI.** In your response, indicate the group and question number you are responding to using the provided format (e.g., I.a, II.b). Appropriate citations are highly encouraged. Respondents are also welcome to address other relevant avenues or technologies that are not outlined below, except for those that fall under the "Areas Not of Interest" described above.



I. Relevant, High-Impact Catalysis for Net Zero

Because of the vast number of applications for catalysts in the renewable energy, emission reduction, and CO₂ conversion space, ARPA-E is seeking insights into where an accelerated workflow would bring about the greatest impact in efficiency, emissions reductions, and/or precious metal reductions. ARPA-E is restricting responses to only include heterogeneous catalysis for electrochemical and thermochemical reactions. It would be beneficial to select reference chemistries from each class so that the workflow and knowledge obtained can translate to other catalytic processes of the same type. For example, determining if technology developed for optimizing the workflow from material to device for the oxygen evolution reaction could be used for the hydrogen evolution reaction would be of interest. Further, reference chemistries may serve as calibration tools within a cohort, especially if commercial scale data is available. Barriers to adoption, such as infrastructural and economic barriers, are also points of interest.

Section I Questions: Relevant, High-Impact Catalysis for Net Zero

- I.a. What are key electrochemical or thermochemical processes that will be required to reach the U.S. 2050 net zero emissions goal? What is the estimated energy requirement of these processes and how does the catalyst/sorbent affect these energy requirements?
- I.b. Given the answer in I.a., what are similar processes or reactions that would benefit from the optimization of I.a.? How do these affect future U.S. energy requirements?

II. Hardware for Accelerating and Scaling High Impact Catalysis

High-throughput screening and synthesis workflows have been widely deployed in the biotechnology industry for over three decades.²⁸ The success in developing these tools benefited from the significant capital investment of biotechnology R&D organizations, as well as the inherently easier nature of handling liquid-phase synthesis of small molecules within parallel testing environments. Active heterogeneous catalysts, which are often in the form of nanoparticles, are much more difficult to synthesize and handle, leading to less deployment onto high throughput synthetic systems and platforms.²⁹ Recently, guided automation has been proposed as a different, but clearly synergistic route compared to high throughput parallel synthesis. These tools can be lower in capital than traditional high throughput equipment, built upon modularity, and are designed to be compatible with a “big data” approach.³⁰ Whereas conventional high throughput methods are driven more by the speed and combinatorial approaches, newer guided automation approaches are driven more by the combination of hardware and software guidance.

²⁸ D A Pereira and J A Williams, “Origin and Evolution of High Throughput Screening,” *British Journal of Pharmacology* 152, no. 1 (September 2007): 53–61.

²⁹ Howard W. Turner, Anthony F. Volpe, and W.H. Weinberg, “High-Throughput Heterogeneous Catalyst Research,” *Surface Science* 603, no. 10–12 (June 2009): 1763–69.

³⁰ John M. Gregoire, Lan Zhou, and Joel A. Haber, “Combinatorial Synthesis for AI-Driven Materials Discovery,” *Nature Synthesis* 2, no. 6 (March 9, 2023): 493–504.



Specifically, there are three major areas where the hardware of high-throughput, automated, or parallel approaches are of interest:

1. **Catalyst synthesis:** Typically, the variation in catalyst synthesis occurs from exploring the compositional space. This may be through different stoichiometric ratios of precursor materials to achieve compositional variance of the catalyst particle. Validation of desired composition may use a combination of elemental and crystal structural analysis tools.
2. **Composite structuring:** The nanostructure of the catalyst and structure of the composite material expands the parameter space further. Each composition can have an unknown number of stable nanostructures with different catalytic activity that is then altered by compositing, machining, or engineering to form distinct microstructures. Handling the fabrication processes and characterization requires additional hardware innovation.
3. **Device testing:** The high throughput fabrication of composite catalytic materials, new processing condition discovery, as well as device characterization (including duration testing) further expands the parameter space. Due to the number of additional components and complexity involved in testing at the device-level, major bottlenecks and longest timescales are expected in this level.

For each of the following Section II questions, indicate the target application/reaction, type of reaction, or any other relevant details.

Section II Questions (Part 1): Systems for Accelerating Neat, Nanostructured, Active Material Synthesis and Characterization

- II.a. What is the state of the art for synthesizing catalytic materials based on compositional variance? What are the limitations of compositional synthesis presently performed?
- II.b. For synthetic procedures done in some accelerated manner, how reproducible are the catalytic material syntheses?
- II.c. What are characterization methods that are compatible (either in situ or integrated) with high throughput material synthesis? What are the limitations of these methods?
- II.d. What are the major barriers to adoption for platforms that accelerate compositional degrees of freedom in catalyst synthesis?

Section II Questions (Part 2): Systems for Accelerating Composite, Microstructural Catalysts Formulated and Compounded from Active, Neat Materials

- II.e. What advancements would need to be integrated into accelerated compositional varying platforms to control the material nanostructure and up to composite microstructure?
- II.f. What is the timescale duration that may be increased due to addition of the structural degree of freedom?



Section II Questions (Part 3): Reaction Systems for Accelerating the Testing of Engineered Forms of Catalytic Composites for Scale-Up

- II.g. What engineered form of catalytic material (e.g., monolith, pellet, MEA, contactor) can likely be correlated to a scaled-up reaction device?
- II.h. What advancements would be needed to integrate this level of device testing with accelerated compositional variance testing described in Section II (Part 1)?
- II.i. Are there identified proxies or empirical relationships in the reaction of interest that can give guidance or down select from Section II (Part 1 or Part 2) to determine the most promising materials for device-level integration and testing?

III. Experimental Data Science, ML, and AI Applied in Catalysis Development Cycles

Data science, ML, and AI have transformed the way researchers can access and analyze large amounts of multi-dimensional data to draw faster conclusions than a traditional, serial approach lacking these capabilities. Data science and ML can efficiently analyze large amounts of multi-dimensional data, predict outputs of interest, and enable accelerated data-driven workflows. Data generated from experimental testing enhances future analyses and ML predictions. For instance, in the example given in Section II, an accelerated hardware platform can generate a significantly large number of compositionally different catalytic material compositions, structures, and ultimate device parameters. A development workflow which integrates synthesis of the catalytic materials with “testable” devices may result in a multi-dimensional parameter space that necessitates state-of-the-art data-driven approaches.

Section III Questions:

- III.a. What automation techniques exist to streamline the way experimental catalysis data is acquired, managed, and shared internally with a team?
- III.b. What key variables and decisions need to be considered for applying statistical, data science, and ML tools to catalytic application testing?
- III.c. What statistical, ML, and AI tools exist for processing experimental catalysis data, as well as discovering and optimizing catalytic materials? What are the challenges?
- III.d. How are “experts in the loop” best utilized to guide and refine the data-driven development of catalysis?

IV. Coupled Multi-Scale Modeling and Accelerated Experimentation in Catalytic Development Cycles

Each stage of the catalytic development cycle can utilize modeling to explain or determine performance, ranging from atomistic *ab initio* calculations to device-scale multiphysics modeling, and even beyond to process-level techno-economic analyses. However, many of these efforts are not integrated together, or are disparate from experimental results.

In the accelerated workflow utilizing hardware and software described in Sections II and III, there is opportunity for simulations and modeling to elucidate fundamental mechanisms that may be difficult to uncover in siloed workflows, especially at the composite level. This effort may have synergy with some



forms of existing data repositories in projects, such as the Materials Project and Materials Genome.^{31,32} As the hierarchy increases from catalytic material-to-device, the physics needed to explain the phenomena is distinct at different levels. For instance, at the catalytic material level, *ab initio* calculations and molecular dynamic simulations may be the most relevant. However, device level performance may require multiphysics including fluid dynamics. With the advances described in Section II, it may be possible to integrate or synthesize the knowledge into models that are currently being developed in a siloed fashion. In addition, data science tools may accelerate the modeling of specific simulations.

Section IV Questions:

- IV.a. Which modeling processes serve as a bottleneck in the proposed workflow concept and may benefit from acceleration? How does this acceleration lead to useful advancements in understanding the fundamental problems?
- IV.b. What are the difficulties in integrating modeling techniques across different length scales? Can this be accomplished/handled with data science approaches?

³¹ "Materials Project - Home," Materials Project, <https://next-gen.materialsproject.org/>.

³² "Materials Genome Initiative | WWW.MGI.GOV," <https://www.mgi.gov/>.