A catalyst acceleration platform towards realizing the energy transition

Marcus Tze-Kiat Ng†, Ahmed S. M. Ismail†, Alexander J. S. Hammer†*

Dunia Innovations, Rheinsberger Str.76/77, 10115 Berlin, Germany

† Authors contributed equally

**Correspondence: alex@dunia.ai*

Summary

The climate emergency has made it necessary to rethink our economy which relies heavily on fossil fuels. Currently, sustainable processes are too costly, but catalysis as a key enabling technology has the potential to reduce process costs to a level that makes them economically viable. In spite of that, the existing catalyst discovery paradigm depends heavily on serendipity and trial and error methods. What is urgently needed to transform the energy transition is a catalyst acceleration platform (CAP) that expedites the development of next-generation sustainable processes. To advance the field, we need to use state-of-the-art robotic and algorithmic tools to look beyond the well-established systems that dominate the current research landscape. Herein, we discuss the requirements for a successful catalyst acceleration concept and the societal impact of breakthrough catalytic materials.

Introduction

Our world is about to undergo a great reallocation of resources.[1] Deployment of current technologies hold promise to lead to massive reductions in greenhouse gas emissions but to reach the climate targets established in the Paris agreement, we need to decarbonize all industrial processes including those in hard-to-abate sectors such as chemicals or aviation (Figure 1). A range of innovative technologies such as carbon upcycling was proposed to achieve the transformation to a circular carbon economy. However, many processes remain far from deployment due to prohibitive operating costs.

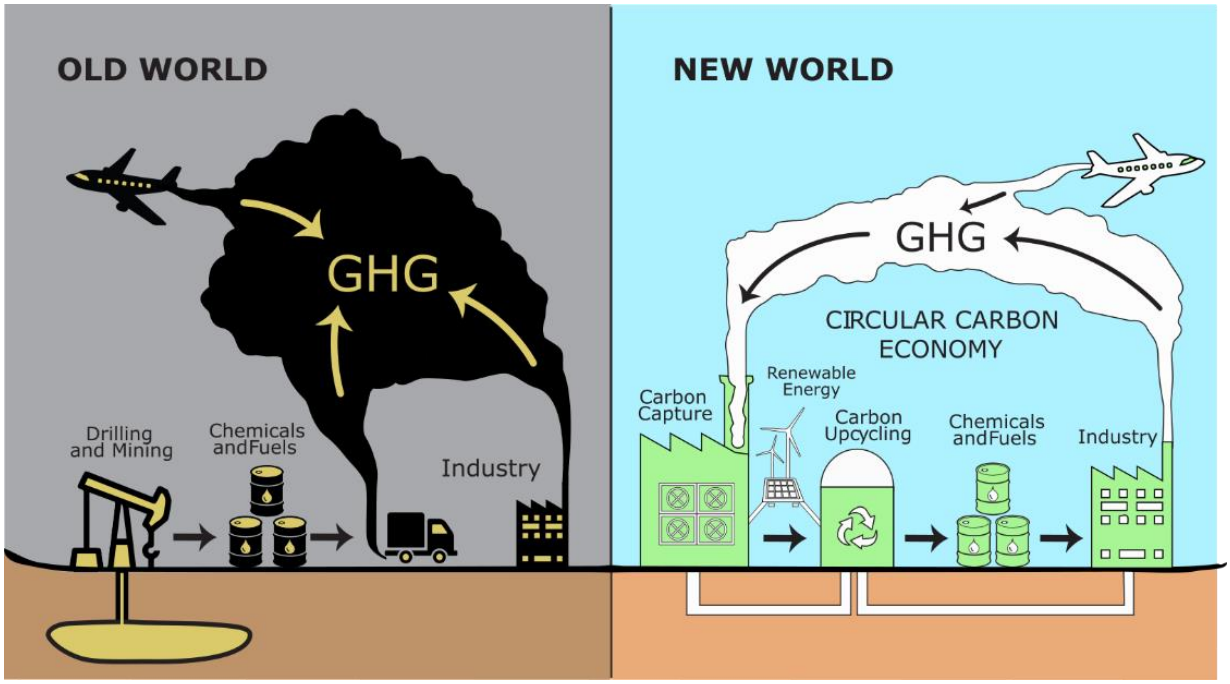

Figure 1. The transition from a fossil fuel-based economy to a circular economy requires a large reallocation of resources and deployment of new clean processes and technologies. *Left: Today's world relies on fossil fuels for energy, transportation, and industries that provide downstream consumer goods such as textiles or food. Right: In tomorrow's world, emissions are avoided, reduced or negated through biomass, carbon capture, and direct air capture technologies. Carbon-neutral fuels such as hydrogen are used to power large parts of the industry, and consumer goods are supplied via carbon-neutral or negative pathways.*

Status Quo in Catalysis

In our current economy, catalysts drive 90% of chemical processes and therewith impact 30-40% of the world's GDP.[2,3] They are key to lowering materials consumption and energy requirements but often take decades of research and successful candidates are often discovered serendipitously. Following this line of thinking, the development of the next generation of catalytically active materials appears to be a priority for clean technology research and development (Figure 2).

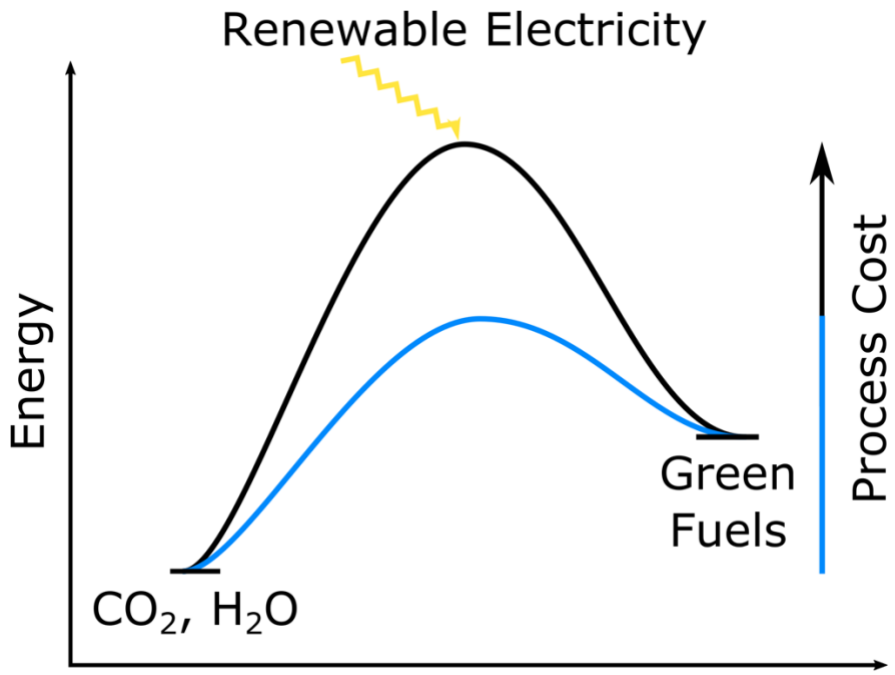

Figure 2. Catalysis is an enabling technology for the provision of carbon-neutral fuels. *Black line: The renewable energy required to produce competitive carbon-neutral fuels is too high leading to too high process costs. Blue line: Efficient catalysts capable of converting inert molecules such as water and carbon dioxide into fuels and bulk chemicals could lower the operating cost of such processes significantly.*

Automation is crucial in catalysis due to advantages such as increased safety, scalability and decreased long-term cost.[4] However, most automation tools available commercially are at most suitable for homogeneous catalysis, which only accounts for 15% of catalytic processes.[5] This is because homogeneous catalysts are typically either dissolved in liquid or come as a suspension, which simplifies the automated workflow. Heterogeneous catalysis, on the other hand, has more challenges when it comes to automation. Whilst today's catalytic processes are based on thermal catalysts, the future of catalysts will evolve around electrocatalysts as the world moves towards electrification of the industry through electricity/energy generated from renewable sources.[6] To date, there is no automated commercial apparatus that is designed to deal with synthesis and testing of electrocatalysts in the market, as it is an emerging field compared to the traditional and more established thermal catalysts. While thermocatalyst discovery could and should be accelerated as well, the focus of the remainder of this paper is placed on electrocatalysis. As a

rather nascent field it would greatly benefit from accelerated materials discovery and that plays an increasingly important role in the energy transition which will be fueled by renewable electricity instead of fossil fuel combustion.

Computational atomistic models have been utilized by scientists to rationalize the experimental phenomena of the catalysts. Traditional *ab initio* simulation methods, such as Density Functional Theory (DFT), remain the main tools to obtain atomistic insight into intriguing experimental results.[7] One of the platforms available is Catalysis-Hub, a database of surface reactions generated with DFT.[8] Researchers around the world have spent a considerable amount of time and resources to simulate optimal electrocatalysts for various processes (e.g., hydrogen evolution reaction, oxygen evolution reaction, hydrogen oxidation reaction, oxygen reduction reaction, and carbon dioxide conversion to various products).[9,10] However, despite their importance in understanding catalytic processes on the atomic scale, these methods are limited by their computational complexity, which often renders them more useful for post-rationalization rather than for generating novel materials hypotheses.[11] All computationally derived datasets have the same bias as the computational methods in terms of classes of materials that can be correctly characterized or in terms of sampling information, often limited to structural analysis with few insights in the catalyzed reactions in which the catalytic species might be involved.[12] More importantly, the catalysts designed purely in-silico do not provide synthetic routes and often a team of experts are required to find potential pathway to synthesize. Meanwhile, experimental datasets are available through literature, but pre-processing procedures may be unclear and hard to replicate. Often, these datasets are incomplete in the information they provide and representative only of a restricted subset of molecules or reactions.

Recent advances in computation allow the exploitation of artificial intelligence to accelerate the development of new catalysts.[13] For example, machine learning has the potential to provide quantum chemical accuracy at force field computational cost.[14] Furthermore, AI can make substantial contributions to catalysis in terms of direct generation of novel catalyst hypotheses, provided appropriate data, including storing and representation strategies are available.[15,16] In contrast to protein structure prediction, large high-quality datasets are available in the public domain[17,18] and AI approaches such as AlphaFold[19] disrupted the state-of-the-art, emerging fields

such as electrocatalysis lack high-quality experimental databases, which severely hinders the benefits of data-driven discovery approaches.[20] A notable example of leveraging AI in conjunction with more easily accessible simulation data is the Open Catalysts Project 2020 (OC20), which provides 1,281,040 structural relaxation trajectories calculated with DFT across a wide range of materials which was extended further.[21-23] Solid state material databases exist but mainly relate the catalyst performance only to its bulk properties.[24,25] Overall, the application of machine learning holds great promises to accelerate the discovery in underexplored fields such as bio-[26] and electrocatalysis[16]. At the same time, lack of standardized testing and poor reproducibility leads to poor data quality that cannot be leveraged in algorithmic searches.[27-29]

Catalysis Acceleration Platform (CAP)

To discover breakthrough catalytic materials, we believe that we need to debias the search away from well-established foundations and boldly explore the depth of chemical space (Figure 3).[30] While both expert-guided experimentation and early MAP implementations typically focus on rather narrow areas of chemical space, it is our expectation that finding answers to societal solutions will require a materials acceleration concept that enables global exploration combined with local exploitation of the acquired information. Many pioneering works highlight the potential of automated and data-driven workflows in catalysis. At the same time, a high-performing but unstable catalyst is not scalable and the urgent need to discover novel, scalable catalysts to realize the energy transition remains unchanged. Therefore, we must scale up our efforts to unlock radical innovation.

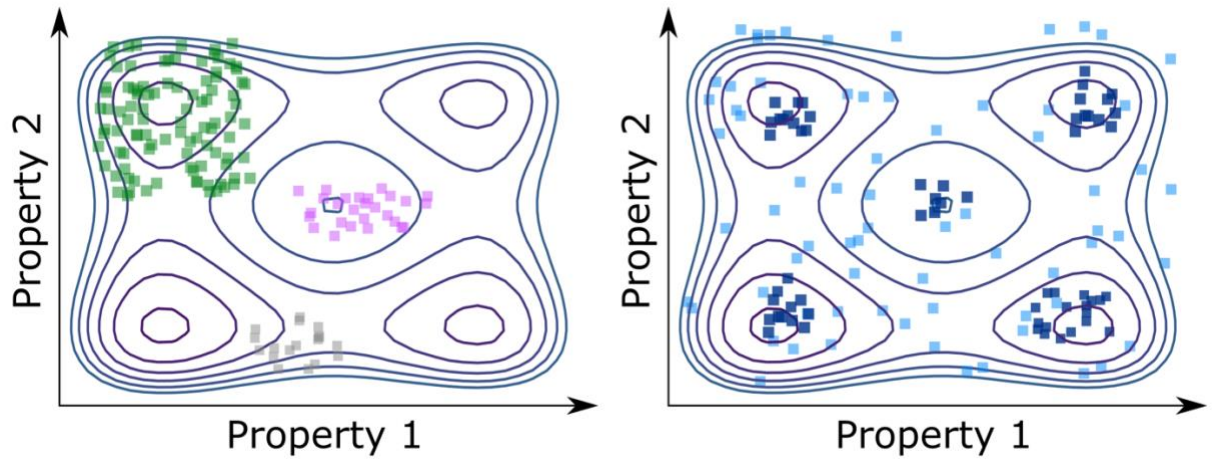

Figure 3. An illustration of bias in both traditional R&D and early MAP implementations compared to a bold materials acceleration concept for societal solutions in an example two-dimensional search space. *Left: Green, purple and grey points represent experiments suggested by researchers based on experience, intuition and literature precedent. This often leads to narrow constraints in which local optima are identified. Right: An unbiased materials acceleration platform for catalysis that generates large quantities of data and combines global exploration (light blue) with local optimization (dark blue).*

As such, our concept relies on identifying catalyst candidates in a highly combinatorial search space using advanced sample-efficient algorithms (Figure 4). This requires combining expertise from automation, computation and cutting-edge catalysis science into a state-of-the-art robotic platform flexible enough to handle the vast variety of chemical inputs, synthesis, and processing conditions for surface-engineered catalytic materials.[31] Rather than employing massive parallelism and miniaturization, our solution stays as close as possible to industrially relevant synthesis and testing conditions to ensure ready scalability from laboratory to pilot. A scalable platform should further adopt standards in laboratory automation hardware[32] and capture all relevant data and metadata generated throughout a given sample history, including quality control data for anomaly detection and crucial data from failed experiments[33] in a findable, accessible, interoperable and reusable manner.[34,35] This alongside standardized catalytic testing and characterization will enable a higher degree of reproducibility than traditional research approaches.[36] Knowledge-informed machine learning pipelines that allow the incorporation of the results of quantum chemical simulations into the active learning loop will be of high

importance to enable sample-efficient searches.[37-39] Using such materials-aware representations, these initial 'hits' can be further optimized into promising catalyst candidates within a constrained search space focusing on a narrow area. In this process we generate high quality and diverse training data to train both global and local models to predict catalytic performance.

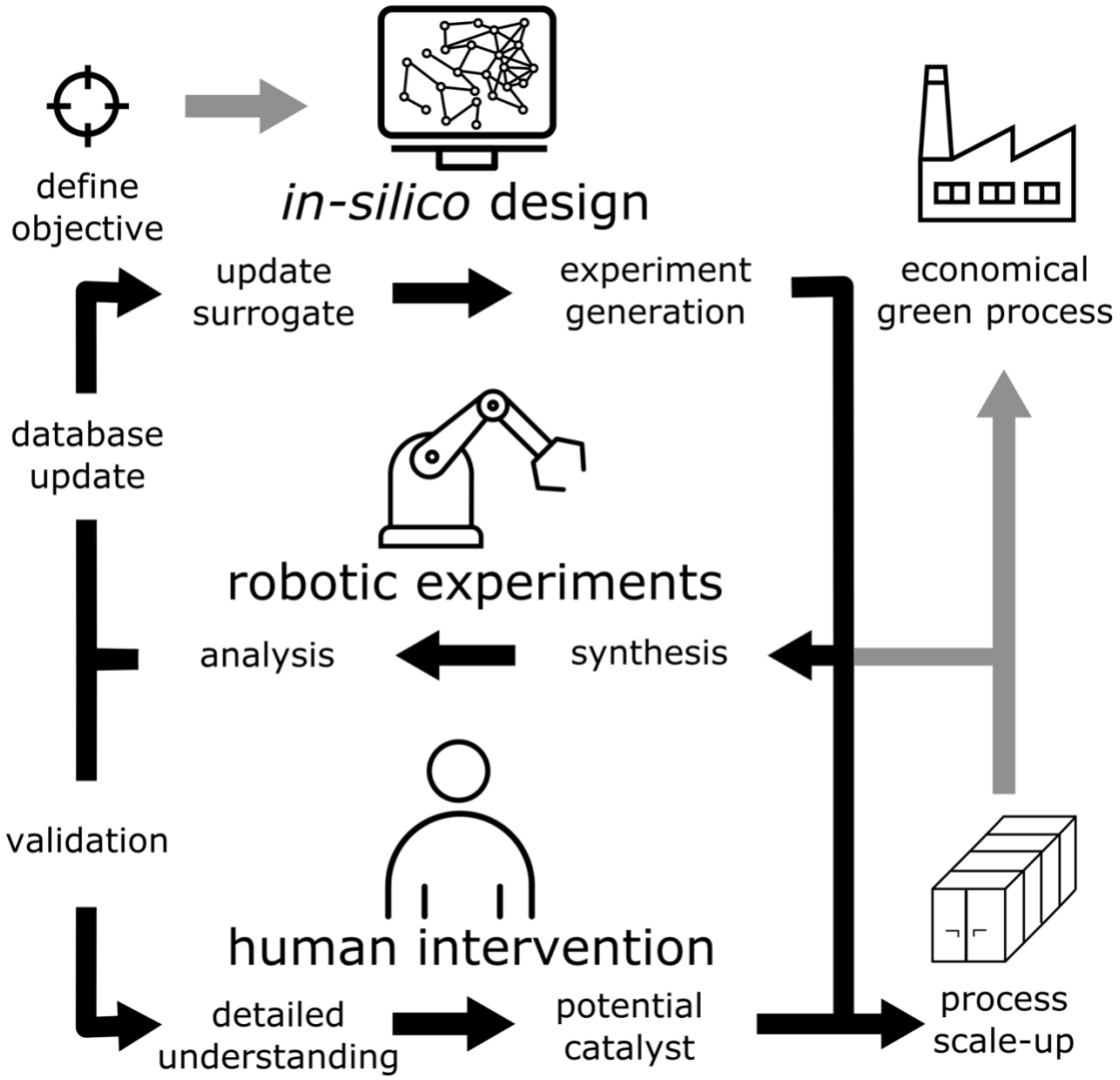

Figure 4. Workflow of the Catalyst Acceleration Platform (CAP). An algorithm selects the experiments to carry out based on chemicophysical properties of the catalyst to achieve a user-defined objective. Experiments are carried out by the robotic platform to ensure production of reproducible high-quality data. The surrogate model updates based on the results and this closed-loop iterative approach is carried out until an interesting result is achieved and requires further validation by a human. The potential catalyst is then scaled up towards the industrially relevant

scale for the given chemical process. Results and metadata of each process are recorded and fed back into the loop.

The challenge is to link the many existing approaches and enable feedback loops to quickly iterate computational design, synthesis, testing and scale-up without data losses. Uniquely, our approach links the full synthetic process conditions to the catalysis performance, focusing not only on activity and selectivity but emphasizing stability, cost and manufacturability which are paramount to translating discoveries into scalable processes. While leveraging the data-driven decision-making paradigm of self-driving labs, we believe that full autonomy at the expense of human creativity can hinder scientific progress. Therefore, an efficient platform must be designed to maximize the productivity of domain experts and provide interfaces for researchers to refine the search space, study unexpected phenomena and immediately translate discoveries into practice. In order to capitalize on the promised benefits of current approaches, novel solutions in materials-aware machine learning algorithms, consistent data ontologies, and flexible automated workflows are required. Comparably to efforts in the AI-driven COVID-19 drug discovery space, generated datasets could be released to the public to stimulate innovation, inviting the community to contribute to the next generation of machine learning models.[40,41] In turn, proprietary models could be released in a similar fashion as IBM RXN to benefit the wider community without compromising commercial interests.[42] Such efforts alongside breakthrough discoveries would further aid in building trust into the AI-driven science paradigm.

Environmental and Economic Impact of Breakthrough Catalysts

Many sectors can be decarbonized through expanding renewables, electrification, energy efficiency and emission reduction measures. Among the hardest-to-abate sectors Chemicals, Aviation and Shipping account for around 9.4% of global greenhouse gas emissions and their relative share is bound to increase significantly as we decarbonize the easier-to-abate sectors.[43] In this section, we examine how discovery, scale-up and commercialization of CO_2-to-X catalysts with >60% CO_2 conversion efficiency could aid in achieving a net zero carbon economy. [43] Of

course, many new sustainable processes are required to fully realize the energy transition including hydrogen generation, electrochemical ammonia synthesis and splitting are among them, with catalysis being a key enabling technology in each case. [44]From an environmental point of view, capturing CO_2 from air or from biomass and converting it to one of the base chemicals such as methanol would reduce CO_2 emissions by up to 95% and nitrogen oxide by up to 80%.[45] The reduction of these two greenhouse gases amounts to a reduction of 0.3% of the global greenhouse gas emissions that are normally emitted from methanol production from fossil-based feedstock.[46] Similarly, enabling the catalytic conversion of CO_2 to green ethylene with a CO_2 conversion efficiency that is higher than 60% would remove 260 million tons of CO_2 emissions per year, which is around 0.8% of the global annual CO_2 emissions.[47] These examples demonstrate the tremendous environmental benefits of introducing sustainable catalysts that effectively convert an inert molecule such as CO_2 to green fuels and chemicals.

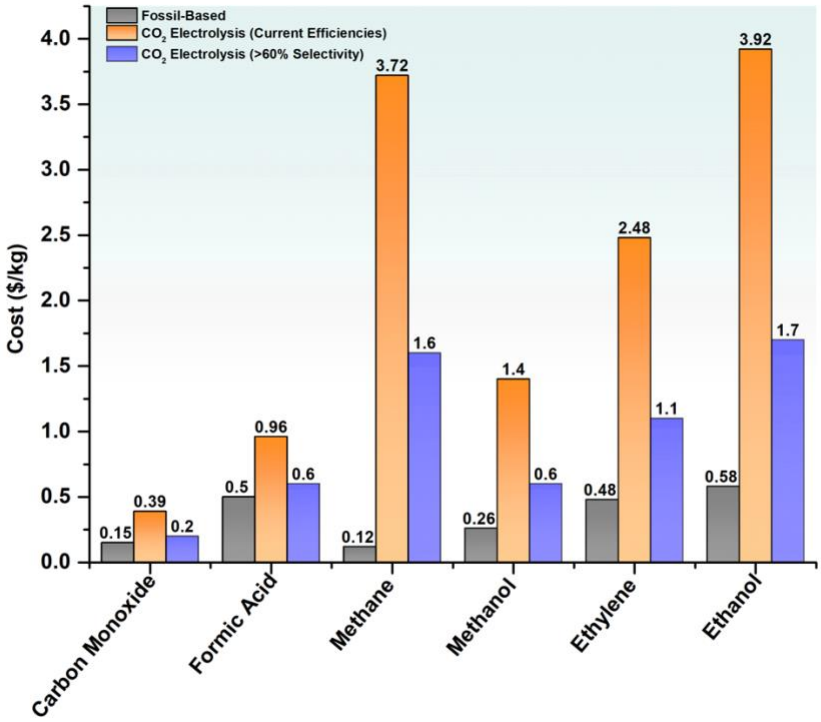

Figure 5. The economics of CO_2-based chemicals. The effect of the electrochemical CO_2 conversion efficiency on the cost reduction of CO_2-based products relative to fossil-based chemicals. The analysis is based on the data produced by the authors and data compiled from the literature.[44,48,49]

From an economic perspective, discovering a highly active and stable CO_2 conversion catalyst could unlock significant economic benefits as there are at least 50 different CO_2-to-chemical approaches that could be realized by 2030 based on the authors' own analysis. For example, the global methanol market is expected to grow from USD 24.10 billion in 2020 to USD 66.06 billion by 2030 at a compound annual growth rate (CAGR) of 5.53%, and the increase in demand is being driven predominantly by adopting green methanol as a marine fuel to reduce the carbon footprint in marine shipping. Similarly, the highly demanded base chemical ethylene has an overall market size of USD 113 billion in 2021 and is expected to grow to USD 125.02 billion in 2022 at a compound annual growth rate (CAGR) of 10.6%. Introducing novel catalysts that can enable highly selective electrochemical conversion of CO_2 to base chemicals and fuels could reduce the operational expenses by up to 32% or by around USD 331 million per electrochemical CO_2 conversion process within a 5-year period (Figure 5).[44,48] This demonstrates the importance of catalysts in driving a transition to more sustainable chemical processes.

Furthermore, the MAP-based catalyst discovery approach offers several advantages in terms of experimentation cost and time. Based on our analysis, building such a platform and carrying out a catalysts discovery campaign for 2-3 years will cost 70% less than carrying a similar campaign following the traditional model where several labs with complementary expertise collaborate on a catalysis discovery project. This cost-effective method will not only facilitate the discovery of novel catalysts, but it can also enable global equitable access to novel technologies for materials discovery. This is in addition to the better data reproducibility since the main catalyst synthesis data and the experimental metadata will be stored in a central database for future repetitions of the experiments if needed. Finally, it is estimated that implementing this accelerated materials discovery approach could shorten the time-to-discovery by up to 90%, which will open several new materials discovery avenues and will create new opportunities for future generations to work on solutions that will enable the sustainable energy transition to proceed in a faster pace towards global deployment.[50]

Ultimately, the catalyst acceleration platform could lead to novel catalytic materials with the potential to enable sustainable chemistry processes that are more efficient, selective, and cost-effective. Success would indeed solve societal problems by enabling CO_2 emission reduction on

a gigaton scale and providing new economic opportunities around the globe. To achieve this vision, significant R&D funding will be needed to develop and run the catalyst acceleration platform to explore chemical space on an unprecedented scale. More importantly, commercial partners who are willing to engage in green processes to enhance the robustness of catalysts and explore their potential for expansion are required. If humanity decided to engineer coffee machines and weaponry to such a sophisticated level, we should allocate appropriate resources to accelerate for the development of green processes that would otherwise come too late to achieve our climate targets.

Acknowledgments

We gratefully acknowledge support from the Advanced Materials Lab, ESF Project Number: 2022000024, Innovation Network for Advanced Materials e.V, IRIS Adlershof Institute and the Humboldt-Innovation GmbH of Humboldt University of Berlin, Helmholtz Innovation Lab HySPRINT of Helmholtz-Zentrum Berlin, and the Berlin Senate.

Author Contribution

All authors contributed equally to this work.

Declaration of Interest

The authors are the founders of Dunia Innovations UG, a company working on accelerated catalyst development for sustainable processes.

References

1. Krishnan, M., Samandari, H., Woetzel, J., Smit, S., Pacthod, D., Pinner, D., Nauclér, T., Tai, H., Farr, A., Wu, W., and Imperato, D. (2022). The net-zero transition: What it would cost, what it could bring. McKinsey & Company.
2. Catlow, C.R., Davidson, M., Hardacre, C., and Hutchings, G.J. (2016). Catalysis making the world a better place. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences *374*, 20150089. https://doi.org/10.1098/rsta.2015.0089.

3. Hagen, J. (2015). Economic Importance of Catalysts. In Industrial Catalysis, pp. 459-462. https://doi.org/10.1002/9783527684625.ch17.
4. Trobe, M., and Burke, M.D. (2018). The Molecular Industrial Revolution: Automated Synthesis of Small Molecules. Angewandte Chemie International Edition *57*, 4192-4214. https://doi.org/10.1002/anie.201710482.
5. dos Passos Gomes, G., Pollice, R., and Aspuru-Guzik, A. (2021). Navigating through the Maze of Homogeneous Catalyst Design with Machine Learning. Trends in Chemistry *3*, 96-110. https://doi.org/10.1016/j.trechm.2020.12.006.
6. Koshy, D.M., Nathan, S.S., Asundi, A.S., Abdellah, A.M., Dull, S.M., Cullen, D.A., Higgins, D., Bao, Z., Bent, S.F., and Jaramillo, T.F. (2021). Bridging Thermal Catalysis and Electrocatalysis: Catalyzing CO2 Conversion with Carbon-Based Materials. Angewandte Chemie International Edition *60*, 17472-17480. https://doi.org/10.1002/anie.202101326.
7. Dickens, C.F., Kirk, C., and Nørskov, J.K. (2019). Insights into the Electrochemical Oxygen Evolution Reaction with ab Initio Calculations and Microkinetic Modeling: Beyond the Limiting Potential Volcano. The Journal of Physical Chemistry C *123*, 18960-18977. https://doi.org/10.1021/acs.jpcc.9b03830.
8. Winther, K.T., Hoffmann, M.J., Boes, J.R., Mamun, O., Bajdich, M., and Bligaard, T. (2019). Catalysis-Hub.org, an open electronic structure database for surface reactions. Scientific Data *6*, 75. https://doi.org/10.1038/s41597-019-0081-y.
9. Nørskov, J.K., Bligaard, T., Rossmeisl, J., and Christensen, C.H. (2009). Towards the computational design of solid catalysts. Nature Chemistry *1*, 37-46. https://doi.org/10.1038/nchem.121.
10. Seh, Z.W., Kibsgaard, J., Dickens, C.F., Chorkendorff, I., Nørskov, J.K., and Jaramillo, T.F. (2017). Combining theory and experiment in electrocatalysis: Insights into materials design. Science *355*, eaad4998. https://doi.org/10.1126/science.aad4998.
11. Burke, K. (2012). Perspective on density functional theory. The Journal of Chemical Physics *136*, 150901. https://doi.org/10.1063/1.4704546.

12. Jain, A., Shin, Y., and Persson, K.A. (2016). Computational predictions of energy materials using density functional theory. Nature Reviews Materials *1*, 15004. https://doi.org/10.1038/natrevmats.2015.4.

13. Tran, K., and Ulissi, Z.W. (2018). Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution. Nature Catalysis *1*, 696-703. https://doi.org/10.1038/s41929-018-0142-1.

14. Johansson, A., Xie, Y., Owen, C.J., Lim, J.S., Sun, L., Vandermause, J., and Kozinsky, B. (2022). Micron-scale heterogeneous catalysis with Bayesian force fields from first principles and active learning. arXiv:2204.12573. https://doi.org/10.48550/arXiv.2204.12573.

15. Freeze, J.G., Kelly, H.R., and Batista, V.S. (2019). Search for Catalysts by Inverse Design: Artificial Intelligence, Mountain Climbers, and Alchemists. Chemical Reviews *119*, 6595-6612. http://doi.org/10.1021/acs.chemrev.8b00759.

16. Zhong, M., Tran, K., Min, Y., Wang, C., Wang, Z., Dinh, C.-T., De Luna, P., Yu, Z., Rasouli, A.S., Brodersen, P., et al. (2020). Accelerated discovery of CO2 electrocatalysts using active machine learning. Nature *581*, 178-183. https://doi.org/10.1038/s41586-020-2242-8.

17. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Research *28*, 235-242. http://doi.org/10.1093/nar/28.1.235.

18. Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol *112*, 535-542. http://doi.org/10.1016/s0022-2836(77)80200-3.

19. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. http://doi.org/10.1038/s41586-021-03819-2.

20. Trunschke, A., Bellini, G., Boniface, M., Carey, S.J., Dong, J., Erdem, E., Foppa, L., Frandsen, W., Geske, M., Ghiringhelli, L.M., et al. (2020). Towards Experimental Handbooks in Catalysis. Topics in Catalysis *63*, 1683-1699. https://doi.org/10.1007/s11244-020-01380-2.

21. Chanussot, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., Tran, K., Heras-Domingo, J., Ho, C., Hu, W., et al. (2021). Open Catalyst 2020 (OC20) Dataset and Community Challenges. ACS Catalysis *11*, 6059-6072. http://doi.org/10.1021/acscatal.0c04525.

22. Tran, R., Lan, J., Shuaibi, M., Goyal, S., Wood, B.M., Das, A., Heras-Domingo, J., Kolluru, A., Rizvi, A., Shoghi, N., et al. (2022). The Open Catalyst 2022 (OC22) Dataset and Challenges for Oxide Electrocatalysis. arXiv:2206.08917. https://doi.org/10.48550/arXiv.2206.08917.

23. Kolluru, A., Shuaibi, M., Palizhati, A., Shoghi, N., Das, A., Wood, B., Zitnick, C.L., Kitchin, J.R., and Ulissi, Z.W. (2022). Open Challenges in Developing Generalizable Large Scale Machine Learning Models for Catalyst Discovery. arXiv:2206.02005. https://doi.org/10.48550/arXiv.2206.02005.

24. Bergerhoff, G., Hundt, R., Sievers, R., and Brown, I.D. (1983). The inorganic crystal structure data base. Journal of Chemical Information and Computer Sciences *23*, 66-69. http://doi.org/10.1021/ci00038a003.

25. Belsky, A., Hellenbrandt, M., Karen, V.L., and Luksch, P. (2002). New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design. Acta Crystallographica Section B *58*, 364-369. http://doi.org/10.1107/S0108768102006948.

26. Lu, H., Diaz, D.J., Czarnecki, N.J., Zhu, C., Kim, W., Shroff, R., Acosta, D.J., Alexander, B.R., Cole, H.O., Zhang, Y., et al. (2022). Machine learning-aided engineering of hydrolases for PET depolymerization. Nature *604*, 662-667. https://doi.org/10.1038/s41586-022-04599-z.

27. Wang, L., Sofer, Z., and Pumera, M. (2020). Will Any Crap We Put into Graphene Increase Its Electrocatalytic Effect? ACS Nano *14*, 21-25. https://doi.org/10.1021/acsnano.9b00184.

28. Akbashev, A.R. (2022). Electrocatalysis Goes Nuts. ACS Catalysis *12*, 4296-4301. http://doi.org/10.1021/acscatal.2c00123.

29. Keith, J.A., McKone, J.R., Snyder, J.D., and Tang, M.H. (2022). Deeper learning in electrocatalysis: realizing opportunities and addressing challenges. Current Opinion in Chemical Engineering *36*, 100824. https://doi.org/10.1016/j.coche.2022.100824.

30. Jia, X., Lynch, A., Huang, Y., Danielson, M., Lang'at, I., Milder, A., Ruby, A.E., Wang, H., Friedler, S.A., Norquist, A.J., and Schrier, J. (2019). Anthropogenic biases in chemical reaction data hinder exploratory inorganic synthesis. Nature *573*, 251-255. https://doi.org/10.1038/s41586-019-1540-5.

31. Stach, E., DeCost, B., Kusne, A.G., Hattrick-Simpers, J., Brown, K.A., Reyes, K.G., Schrier, J., Billinge, S., Buonassisi, T., Foster, I., et al. (2021). Autonomous experimentation systems for materials development: A community perspective. Matter *4*, 2702-2726. https://doi.org/10.1016/j.matt.2021.06.036.

32. Porr, M., Schwarz, S., Lange, F., Niemeyer, L., Hentrop, T., Marquard, D., Lindner, P., Scheper, T., and Beutel, S. (2020). Bringing IoT to the Lab: SiLA2 and Open-Source-Powered Gateway Module for Integrating Legacy Devices into the Digital Laboratory. HardwareX *8*, e00118. https://doi.org/10.1016/j.ohx.2020.e00118.

33. Raccuglia, P., Elbert, K.C., Adler, P.D.F., Falk, C., Wenny, M.B., Mollo, A., Zeller, M., Friedler, S.A., Schrier, J., and Norquist, A.J. (2016). Machine-learning-assisted materials discovery using failed experiments. Nature *533*, 73-76. https://doi.org/10.1038/nature17439.

34. Nieva de la Hidalga, A., Goodall, J., Anyika, C., Matthews, B., and Catlow, C.R.A. (2022). Designing a data infrastructure for catalysis science aligned to FAIR data principles. Catalysis Communications *162*, 106384. https://doi.org/10.1016/j.catcom.2021.106384.

35. Wulf, C., Beller, M., Boenisch, T., Deutschmann, O., Hanf, S., Kockmann, N., Kraehnert, R., Oezaslan, M., Palkovits, S., Schimmler, S., et al. (2021). A Unified Research Data

Infrastructure for Catalysis Research – Challenges and Concepts. ChemCatChem *13*, 3223-3236. https://doi.org/10.1002/cctc.202001974.

36. Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature *533*, 452-454. http://doi.org/10.1038/533452a.

37. Rueden, L.v., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Walczak, M., Pfrommer, J., Pick, A., et al. (2021). Informed Machine Learning - A Taxonomy and Survey of Integrating Prior Knowledge into Learning Systems. IEEE Transactions on Knowledge and Data Engineering, 1-1. https://doi.org/10.1109/TKDE.2021.3079836.

38. Rohr, B., Stein, H.S., Guevarra, D., Wang, Y., Haber, J.A., Aykol, M., Suram, S.K., and Gregoire, J.M. (2020). Benchmarking the acceleration of materials discovery by sequential learning. Chemical Science *11*, 2696-2706. https://doi.org/10.1039/C9SC05999G.

39. Palizhati, A., Torrisi, S.B., Aykol, M., Suram, S.K., Hummelshøj, J.S., and Montoya, J.H. (2022). Agents for sequential learning using multiple-fidelity data. Scientific Reports *12*, 4694. https://doi.org/10.1038/s41598-022-08413-8.

40. Cuccarese, M.F., Earnshaw, B.A., Heiser, K., Fogelson, B., Davis, C.T., McLean, P.F., Gordon, H.B., Skelly, K.-R., Weathersby, F.L., Rodic, V., et al. (2020). Functional immune mapping with deep-learning enabled phenomics applied to immunomodulatory and COVID-19 drug discovery. bioRxiv, 2020.2008.2002.233064. http://doi.org/10.1101/2020.08.02.233064.

41. Heiser, K., McLean, P.F., Davis, C.T., Fogelson, B., Gordon, H.B., Jacobson, P., Hurst, B., Miller, B., Alfa, R.W., Earnshaw, B.A., et al. (2020). Identification of potential treatments for COVID-19 through artificial intelligence-enabled phenomic analysis of human cells infected with SARS-CoV-2. bioRxiv, 2020.2004.2021.054387. https://doi.org/10.1101/2020.04.21.054387.

42. IBM RXN API. https://github.com/rxn4chemistry/rxn4chemistry.

43. Ritchie, H., Roser, M., and Rablo, R. (2020). CO_2 and Greenhouse Gas Emissions. Our World in Data.

44. Luna, P.D., Hahn, C., Higgins, D., Jaffer, S.A., Jaramillo, T.F., and Sargent, E.H. (2019). What would it take for renewably powered electrosynthesis to displace petrochemical processes? Science *364*, eaav3506. https://doi.org/10.1126/science.aav3506.

45. Hobson, C., and Márquez, C. (2018). Renewable Methanol Report. ATA Markets Intelligence S. L.

46. de Jong, M., Bunse, M., and Hamelinck, C. (2022). Methanol carbon footprint and certification. IMPCA.

47. Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Hauck, J., Olsen, A., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., et al. (2020). Global Carbon Budget 2020. Earth Syst. Sci. Data *12*, 3269-3340. http://doi.org/10.5194/essd-12-3269-2020.

48. Pappijn, C.A.R., Ruitenbeek, M., Reyniers, M.-F., and Van Geem, K.M. (2020). Challenges and Opportunities of Carbon Capture and Utilization: Electrochemical Conversion of CO2 to Ethylene. Frontiers in Energy Research *8*. https://doi.org/10.3389/fenrg.2020.557466.

49. Somoza-Tornos, A., Guerra, O.J., Crow, A.M., Smith, W.A., and Hodge, B.-M. (2021). Process modeling, techno-economic assessment, and life cycle assessment of the electrochemical reduction of CO2: a review. iScience *24*, 102813. https://doi.org/10.1016/j.isci.2021.102813.

50. Seifrid, M., Hattrick-Simpers, J., Aspuru-Guzik, A., Kalil, T., and Cranford, S. (2022). Reaching critical MASS: Crowdsourcing designs for the next generation of materials acceleration platforms. Matter *5*, 1972-1976. https://doi.org/10.1016/j.matt.2022.05.035.