

Data-Driven Innovation in Metal-Organic Frameworks Photocatalysis: Bridging Gaps for CO₂ Capture and Conversion with FAIR Principles

Claudia Bizzarri* and Manuel Tsotsalas*

Metal-organic frameworks (MOFs) have emerged as key materials for carbon capture and conversion, particularly in photocatalytic CO₂ reduction. However, inconsistent reporting of essential parameters in the literature hinders informed decisions about material selection and optimization. This perspective highlights the need for a user-friendly, centralized database supported by automated data extraction using natural language processing tools to streamline comparisons of MOF materials. By consolidating crucial data from scientific literature, such a database promotes efficient decision-making in material selection for CO₂ capture and utilization. Emphasizing the significance of open-source initiatives and the principles of FAIR data—ensuring data are Findable, Accessible, Interoperable, and Reusable—a collaborative approach to data management and sharing is advocated for. Making the database-accessible worldwide enhances data quality and reliability, fostering innovation and progress in CO₂ capture and conversion using MOF materials. Additionally, such databases are valuable in creating artificial intelligence tools to assist researchers in the discovery and synthesis of MOF materials for CO₂ capture and conversion.

1. Introduction

Climate change and global warming are among the most pressing environmental challenges today, raising global


awareness about the need for immediate action.^[1] The increasing levels of carbon dioxide (CO₂), a significant greenhouse gas, must be urgently addressed. Researchers are exploring various technologies to capture, store, and utilize CO₂, transforming it into valuable chemicals that can serve as feedstocks or fuels—thus promoting a circular economy. These technologies, collectively known as carbon capture and utilization (CCU), focus on SEQUESTering CO₂ from industrial emissions or directly from the air for application in diverse processes.^[2] The sustainability of conversion methods is greatly enhanced when powered by renewable energy sources, with sunlight being the cleanest and most abundant option available.^[3]

The development of new and efficient materials that enables the adsorption and photo-driven conversion of CO₂ is fundamental and should take into consideration that CO₂ is usually kinetic and thermodynamic inert.^[4] Natural photosynthesis exemplifies the fixation of CO₂ and solar energy into highly valuable chemicals (e.g., carbohydrates). In artificial photosynthesis, or rather in light-driven CO₂ conversion, the natural process is simplified; specifically, two components are needed: a photosensitizer and a catalyst, which can sometimes be combined into a single material known as a “photocatalyst.” In this perspective, we emphasize the importance to act rapidly in finding economically viable solutions to capture and convert CO₂ into useful products by means of solar energy. To support this goal, accessible and user-friendly databases that consolidate existing knowledge on CO₂ capture and conversion will assist researchers worldwide, enhancing collaboration and the potential to find sustainable answers to the problem of climate change. Moreover, these databases should be capable of extracting important information from reports and articles and be machine-readable to boost efficiency. Researchers are now encouraged to report their data according to the FAIR principles—ensuring that data are Findable, Accessible, Interoperable, and Reusable. The National Science Foundation emphasizes that data management plans should align with these principles to maximize the value of research data.^[5] Similarly, the German Research Foundation (DFG) supports the long-term archiving and open access of research data in line with the FAIR principles.^[6] Moreover, journals such as those published by the Royal Society of Chemistry

C. Bizzarri, M. Tsotsalas
Institute for Organic Chemistry
Karlsruhe Institute of Technology
Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany
E-mail: claudia.bizzarri@uniroma2.it; manuel.tsotsalas@kit.edu

C. Bizzarri
Department of Chemical Science and Technologies
University of Rome “Tor Vergata”
Via della Ricerca Scientifica, 00133 Rome, Italy

M. Tsotsalas
Institute of Functional Interfaces
Karlsruhe Institute of Technology
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aesr.202400325>.

© 2024 The Author(s). Advanced Energy and Sustainability Research published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aesr.202400325

advocate for the implementation of FAIR data principles to advance scientific research.^[7] We envision a new publishing system where, in addition to communicating research through manuscripts, researchers can accurately input numerical data that aid in categorizing and extrapolating information using artificial intelligence^[8] (AI) tools. Machine learning (ML) is a rapidly evolving subfield in AI. By integrating ML tools directly into accessible and machine-readable databases, researchers can more efficiently generate, test, and advance scientific models for CO₂ capture and conversion.^[9]

In this perspective, we discuss the specific requirements of a research data ecosystem focused on metal-organic frameworks (MOFs) as sorbent materials for carbon capture and photocatalytic conversion. A critical issue is the lack of standardized data and measurement protocols. To accelerate the development of new materials for CCU, it is essential to establish unified and reproducible characterization protocols. Developing a comprehensive database will serve as a valuable resource for both academic and industrial researchers, facilitating the sharing of standardized data and comparison of different MOF candidates and process conditions. Additionally, connecting such specialized databases to larger material science repositories like NOMAD^[10] or the Materials Genome Initiative^[11] will enhance data accessibility and promote interdisciplinary collaboration.

2. MOFs in CO₂ Capture and Conversion

MOFs are an important class of crystalline porous materials, built from modular building blocks consisting of metal nodes and organic linkers.^[12] By changing the metals and/or the organic moiety, the physical and chemical properties as well as the geometrical structure of the MOFs can be tailored toward a desired application.

2.1. MOFs for CO₂ Capture

Carbon capture follows two main processes, physisorption (i.e., intermolecular interactions such as van der Waals forces) or chemisorption (i.e., a covalent or ionic bond is formed between the adsorbent material and the CO₂).^[13] In chemisorption, functional groups, usually amines, react reversibly with CO₂ to form carbamides. Upon heating or steam regeneration, the CO₂ is released and the sorbent material regenerated. The MOF material acts in this case as the solid support for the amines, enabling access to the active sites within the porous structure. In physisorption, the CO₂ adsorbs to a solid sorbent material, usually a microporous material. In addition to MOFs, the most common materials are zeolites and porous carbon.^[14] Due to their large design space in composition and structure, MOFs offer the possibility to tailor their properties toward optimal sorbent characteristics, providing an ideal platform to address the challenges related to energy and sustainability.^[15]

2.2. MOF in Photocatalytic CO₂ Reduction

For the last 20 years, MOFs have been actively explored as suitable materials in the heterogeneous photo-driven CO₂ reduction (Figure 1).^[16] The advantage of using MOF instead of

homogeneous systems is related to the higher robustness of the materials and the possibility of separating the (photo)catalyst from the reaction mixture, to be recycled. Moreover, the modularity of construction makes available the incorporation of different units with distinct functionalities. For example, their adjustable porosity not only can be tuned to selectively capture CO₂ but also enhance the catalytic surface area, when compared to other heterogeneous photocatalytic semiconductors. An additional effect of uniform porosity is the scattering of incident light along the internal channels, intensifying the possibility of absorption in the inner core of the MOF.^[17] Likewise, as other semiconductors, their bandgap should be modulated according to the photocatalytic reaction they should promote, together with spatial electronic separation, which could be achieved through charge confinement, avoiding the photogenerated charge recombination. In particular, to activate and reduce CO₂ upon irradiation, the electrons are promoted to the conduction band of the MOF, which is represented by the lowest unoccupied molecular orbital of the metal cluster and should lay above the reduction potential of the proton-assisted CO₂ half-reaction (see Section 3.2). However, to design efficient photocatalysts that are able to exploit solar energy and convert it to chemical energy, the material should absorb in the visible region of electromagnetic radiation. Therefore, the bandgap should be lower than 3.1 eV (corresponding to a wavelength $\lambda > 400$ nm).

The spatial separation and modulation of the properties is possible thanks to the modification of the metal oxo-clusters and the organic linkers, and the enormous amount of research done in this respect demonstrates the customizability of these materials. In addition to their potential as semiconductor photocatalysts, MOFs offer an attractive platform to anchor supramolecular structures, promoting stabilization of the molecular (photo) catalytic species, without decreasing their efficiency, they would have in homogeneous systems. These functionalized materials are usually obtained by modifying the organic linker, that bears a chelating unit, in addition to the carboxylate groups used to coordinate the metal nodes of the MOF. In this way, the active metal center can be incorporated on the MOF backbone. Interestingly, the same modified organic linker can be the ligand for diverse metal centers, acting either as a photosensitizer (enhancing the light-antenna effect of the final functional material), or as a catalyst. Examples of such organic linker are 1,10-phenanthroline dibenzoate^[18] and 2,2'-bipyridine-dibenzoate.^[19] A further example is shown by tetrakis(4-carboxyphenyl) porphyrin that can be linked or grafted to MOF structures through the carboxylate groups, and at the same time, the macrocycle can coordinate different transition metals, such as Fe,^[20] Ni,^[21] Cu,^[22] or Zn,^[23] generating stable and efficient photocatalysts.

Nevertheless, a systematic approach, supported by automated data extraction and ML, is essential to provide straightforward comparison of MOF materials and successfully targeting efficient and sustainable photo-induced processes. Researchers nowadays rely considerably on recent or less recent reviews, to have a valuable overview of the state of the art of a certain topic. Many worthwhile reviews on photocatalytic CO₂ reduction by MOF-based materials have been recently published,^[24] highlighting the considerable interest in this subject. However, a lack of standardized fashion of reporting the results in the original full articles makes the comparison and, thus, the reutilization of data

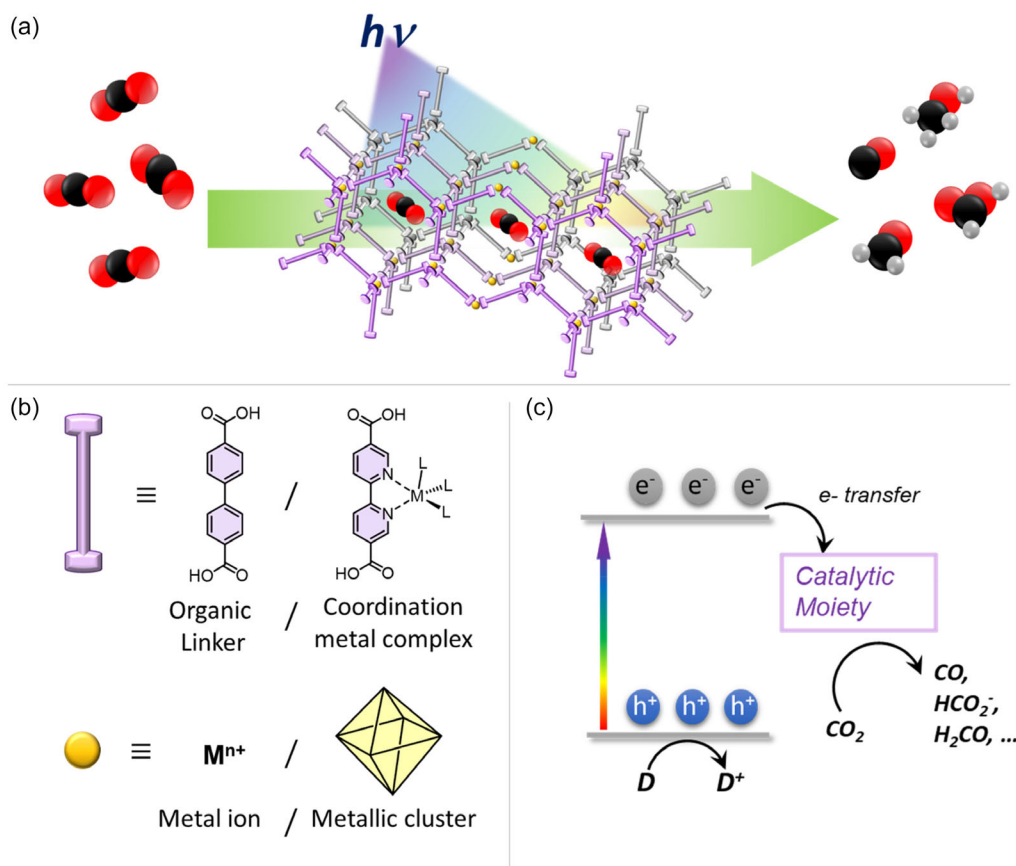


Figure 1. a) Schematic representation of the possible light-induced CO₂ reduction that may happen in MOFs. b) Examples of possible linkers and metallic parts of the MOF. c) Simplified diagram to visualize the major processes involved upon photoexcitation: after light absorption, electron–hole pairs are generated and the excited MOF undergoes an electron transfer, reducing a catalytic moiety (which could be internal or external) that will reduce CO₂. A sacrificial electron donor (D) will restore the ground state of the photocatalyst.

extremely difficult.^[25] In fact, optimization of the MOF structures is only one of the main parameters that should be considered for a certain goal (e.g., photocatalytic CO₂ reduction).

3. Challenges in MOF Research for CCU Applications

3.1. Lack of Standardized Data and Measurement Protocols

In the case of carbon capture, the description of MOF sorbent properties in literature is not well standardized.^[26] Therefore, direct comparison of different MOF materials for carbon capture is challenging. In addition, the relevant performance metrics for carbon capture highly depend on the composition of the carbon source and the carbon capture systems including pre-combustion, oxy-fuel combustion, post-combustion, and direct air capture (DAC) (e.g., the concentration of CO₂ streams in point sources is orders of magnitude higher than in ambient air for DAC). In addition, the high regeneration temperature, water content, and reactive components in the gas mixture of CO₂ pose additional constraints to the material that need to be considered. Synthesizability, cost, and availability of the starting materials

also need to be considered. The adsorption process involves multiple parameters, necessitating the fine tuning of the exact process parameters with the material properties toward the exact CO₂ source.^[27] Even when only considering one process (e.g., the DAC process), the environmental conditions such as humidity and temperature can vary strongly depending on the location of the DAC plant. The exact properties of the ideal sorbent material are therefore strongly process dependent and cannot be generalized.

Readers seeking a detailed discussion on CO₂ capture materials, including key characteristics such as adsorption capacity, adsorption isotherms, selectivity, and stability, are encouraged to consult refs. [28]. We emphasize that the development of standardized measurement and reporting protocols should be a collaborative effort within the research community. Consensus-driven standards are essential for advancing carbon capture research and ensuring the reliability and comparability of results. In recent literature, several reporting guidelines can be identified to make sorbent research more reproducible and comparable across different research labs.^[29]

Previous studies have suggested several key measurements for evaluating new sorption materials for carbon capture applications. CO₂ adsorption capacity, representing the maximum

amount of CO₂ a material can adsorb under specific conditions, has been highlighted as a fundamental metric. Additionally, adsorption isotherms, particularly in the low-pressure range, have been emphasized for their relevance in capturing CO₂ from atmospheric sources or flue gases. The selectivity of materials, particularly their preference for adsorbing CO₂ over other gases such as nitrogen and oxygen, has also been frequently noted. Kinetic studies, which assess the rate of adsorption and desorption, have been recommended as important for determining the efficiency of the process. The literature also stresses the importance of thermal stability, as it indicates the ability of a material to withstand repeated thermal cycles during regeneration. Physical stability and mechanical robustness under operational conditions have been similarly emphasized, alongside chemical stability, which ensures resistance to impurities commonly present in flue gases that could degrade the material. The influence of water vapor on sorbent performance has been recognized as a crucial factor in maintaining CO₂ capture efficiency. Working capacity, defined as the difference between the adsorption and desorption under operational conditions, and cyclic stability, which assesses the material's performance over repeated adsorption-desorption cycles, have been consistently recommended as critical indicators of long-term performance and reusability in these studies.^[30]

In addition to such general guidelines, targeted measurement protocols toward specific applications of carbon capture and specific material classes need to be defined. For example, the following characterization methods were suggested by J. Young et al.^[31] for researchers in material science targeting DAC applications using amine-functionalized sorbents: 1) half-time of CO₂ adsorption at 0.4 mbar of CO₂ and 25 °C; 2) CO₂ adsorption isotherm at 25 °C with a focus on the Henry regime; 3) H₂O adsorption isotherm at 25 °C; 4) density; 5) thermal conductivity; and 6) cyclic stability including the presence of O₂.

3.2. Complexities in Photocatalytic Processes

A promising way to decrease the amount of CO₂ in the atmosphere and concomitantly addressing the exploitation of more sustainable energy sources is the photocatalytic CO₂ reduction. To achieve CO₂ reduction, the (photo)catalyst has to overcome high energetic barriers, due to the intrinsic inertness of this molecule. The basic reduction processes are assisted by protons and have more favorable thermodynamics, although various reaction steps are still needed (Table 1). In particular, we can differentiate the multi-proton and multi-electron transfer reactions to produce C1 products, such as CO, HCHO, CH₃OH, and CH₄, and those that lead to C2+ products, like ethanol, ethane, acetic acid, and so on.^[32] Although the standard reduction potentials of such reactions are advantageous, they still suffer from high kinetic barrier.

Mechanistically, the photo-driven process promotes the redox reaction of the catalyst, facilitating the uptake or the loss of electrons because of the formation of the excited state of the photo-active species. More in detail, the (supra)molecular dye upon absorption of a photon populates the lowest excited state, which is long-lived enough to undergo an electron-transfer process. Depending on whether the dye is luminescent, we can establish if a reductive quenching (so the photocatalyst is reduced) or an oxidative quenching (in this case the dye is oxidized) occurs. At

Table 1. Energy requirements for the proton-assisted CO₂RR.

CO ₂ reduction reaction (CO ₂ RR)	Standard reduction potential [V] vs NHE ^{a)}
CO ₂ + e ⁻ → CO ₂ ^{•-}	-1.91
CO ₂ + 2H ⁺ + 2e ⁻ → CO + H ₂ O	-0.53
CO ₂ + 2H ⁺ + 2e ⁻ → HCO ₂ H	-0.61
CO ₂ + 4H ⁺ + 4e ⁻ → HCHO + H ₂ O	-0.48
CO ₂ + 6H ⁺ + 6e ⁻ → CH ₃ OH + H ₂ O	-0.38
CO ₂ + 8H ⁺ + 8e ⁻ → CH ₄ + 2 H ₂ O	-0.24
2CO ₂ + 8H ⁺ + 8e ⁻ → CH ₃ CO ₂ H + 2 H ₂ O	-0.30
2CO ₂ + 12H ⁺ + 12e ⁻ → C ₂ H ₄ + 4 H ₂ O	0.06
2CO ₂ + 12H ⁺ + 12e ⁻ → C ₂ H ₅ OH + 3 H ₂ O	0.08
2CO ₂ + 14H ⁺ + 14e ⁻ → C ₂ H ₆ + 4 H ₂ O	-0.27

^{a)}NHE = normal hydrogen electrode.

the end, the electron is transferred to the catalytically active species that is able to bind CO₂ and fosters the proton-assisted reduction to various possible products, depending on how many electrons the catalyst can handle.

That is why the development of (photo)catalysts is essential to reduce the activation energy, increasing the reaction kinetics. Moreover, the use of heterogeneous catalysts may promote the formation of multi-electron-reduced products, since the charge accumulation on the surface would be beneficial instead of the limited reduction species of the single-molecular homogeneous catalysts. Understanding of the relation between (photo)catalysts properties and efficiency will positively influence the strategic design of the next material. Moreover, several factors would affect the outcome in the reaction, such as the morphology and structure of the catalyst, defects and vacancies, and active surface areas.^[32d] Moreover, endeavors should be taken for planning large-scale implementation, developing sustainable catalytic materials based on earth-abundant resources.^[33] Thus, the advancement of a practical solar fuel technology could be easily achieved.

3.3. Performance Parameters for the Assessment of Carbon Capture and Photo-Driven CO₂ Reduction

Surveying the literature on MOFs for photo-driven CO₂ reduction, we encountered some parameters, that should be always reported to facilitate a straightforward comparison among the different functional materials.

At present, there are many studies focusing on engineering MOF materials, according to different combinations of metal nodes and organic linkers, with the final goal to improve photocatalytic performance. Nevertheless, every single component in photocatalysis contributes to overall efficiency. Factors like solvent, temperature, additives, power, and time of the irradiation may greatly influence the final performance as well as the composition (and selectivity) of products. Moreover, in heterogeneous systems, total (photo)catalytic efficiency may be affected by mass transfer, thus diffusion is an

important factor, which depends on inherent properties of the semiconductor, such as adsorption capacity and pore size, and on external conditions, such as the stirring velocity inside the reaction batch or the flow stream. Therefore, for the assessment of efficiency of a certain MOF in regard to carbon capture and utilization by means of light, standardized parameters should be given.^[34]

3.3.1. Productivity

The efficiency of a catalyst is usually defined as the number of product moles per catalyst moles (for homogeneous catalysis) or number of active sites (for heterogeneous catalysis). Classically, this ratio is defined as turnover number (TON), which should not be confused with the turnover frequency (TOF), that is the amount ratio between products moles and catalyst active sites per unit of time.^[35] Nevertheless, the use of TON is sometimes controversial as it does not have the same definition in homogeneous and heterogeneous catalysis; therefore, it can generate confusion.^[36] Nevertheless, the efficiency of catalytic conversion for MOFs is often reported in TOF with $\mu\text{mol g}^{-1} \text{h}^{-1}$ as unit. Although the evaluation of productivity in TOF might be satisfactory in thermal reactions, this value does not take into consideration the differences that might arise from diverse light power sources, providing different numbers of photons, influencing greatly the outcome of the photocatalytic reaction. Thus, an important parameter that is needed for the assessment is the photocatalytic yield.

3.3.2. Quantum Yield

A reaction is defined as photocatalytic when the catalytic substance can facilitate the reaction upon exposure to light. Light activation is therefore necessary to initiate and/or accelerate the process. It means that the assessment of the efficiency in a photocatalytic reaction should involve the number of absorbed photons used to drive the reaction. In general, we define photocatalytic quantum yield (QY or Φ) as the ratio between the moles number of products and the moles of absorbed photons. Frequently, the apparent quantum yield (AQY) is reported, which is the ratio between the product moles and the total incident photons. Nevertheless, measuring experimentally the incident photon flux requires extreme care, and also the distance at which the photon flux is determined might influence the final value of the AQY.^[37] Quantum efficiency is a critical parameter for evaluating the performance of a catalyst in a photoinduced reaction. Although there is no specific threshold value of quantum yield required for a reaction to be defined as photocatalytic—since the definition is based on the mechanism rather than the efficacy—it is important to emphasize that photocatalytic quantum yield is a crucial measure of system efficiency.^[38] Moreover, Φ should be considered as a function of irradiation time and power.^[39] Reporting this metric is essential for ensuring fair and accurate comparisons among different studies. By standardizing the measurement and reporting of quantum yield, researchers can better assess and compare the effectiveness of various photocatalytic systems.^[40]

3.3.3. Selectivity of the Photo-Driven Reaction

Herein, we define as selectivity of the photocatalytic process the distribution of different products that may arise from competitive catalytic pathways in the same reaction.^[41] Ideally, a (photo) catalyst is designed in a way that produces only one desired product. However, various reaction conditions can affect the output of the same catalyst. For example, the thermodynamics of different CO_2 reduction processes are very similar (see Table 1), so that a concomitant production of two or more different products is possible. In addition to carbonaceous products, it is common that the selectivity toward a desired product gets lowered because of the favored reduction potential of molecular hydrogen.^[42]

4. Developing a Unified Database of MOF Materials for CCU

Advancing CCU technologies depends on the ability to systematically compare and optimize materials like MOFs. However, the absence of standardized measurement protocols and data sharing platforms hinders progress. In this section, we advocate for the development of a unified, FAIR-compliant database that consolidates MOF research data for CCU applications. The complexity of CCU processes requires a harmonized approach to data collection and reporting. Without standardized measurement protocols, comparing the performance of different MOFs becomes challenging, leading to inefficiencies and duplicated efforts. Standardization enables researchers to replicate studies accurately, validate results, and build upon each other's work.

Several initiatives are paving the way toward standardized data sharing in the CCU community. Examples are databases of single and multicomponent isotherms from literature^[43] as well as the AIF format for sorption data.^[44] Integrating such resources into a unified platform can significantly enhance data accessibility and utility. We advocate for a domain-specific database because this approach allows domain experts to tailor the data to the specific needs and requirements of the field. Nevertheless, such domain-specific guidelines should also be compatible with larger material science data initiatives and built upon their resources and expertise, such as the Materials Genome Initiative or the National Research Data Infrastructure (NFDI).^[45] Ideally, such experimental databases are also compatible with computational material science databases, such as COREMOF^[46] MOFX-DB.^[47]

The use of natural language processing (NLP) tools, such as large language models (LLMs) to extract literature results on carbon capture and utilization, can populate the initial database and allow researcher to easier target their investigation to gaps in the research landscape and identify promising targets.^[48] Furthermore, the use of AI techniques can complement the missing gaps, e.g., add properties such as availability or estimate the mechanical, thermal, or chemical stability of the materials, even if not all such properties are measured on each material explicitly^[49] and guide the user during synthesis and characterization.^[50] Early research also showed potential of employing LLMs for predictive chemistry and as autonomous agent in chemistry.^[51]

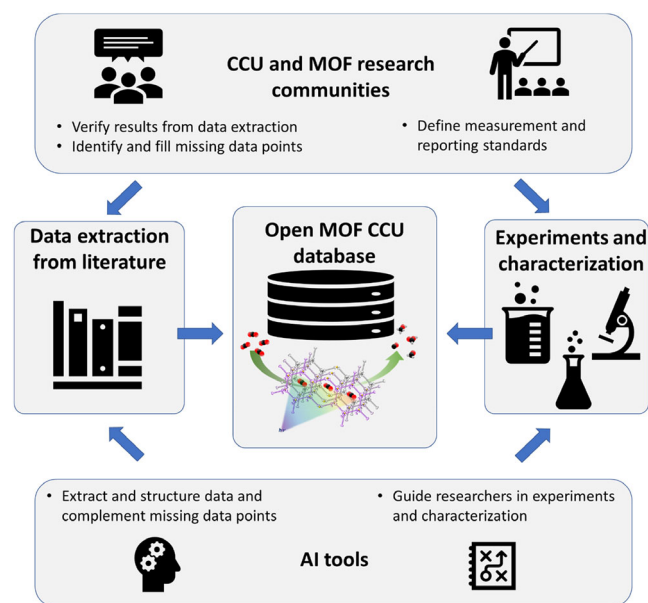


Figure 2. The research data management ecosystem for an open MOF CCU database, designed to adhere to FAIR principles. This ecosystem is composed of several key components: a data repository for storing standardized MOF CCU datasets. The MOF CCU database contains both historic literature data and new datasets, that already are reported in a machine-readable way, following FAIR principles. The MOF and CCU communities play a crucial role to define measurement and reporting standards as well as verifying results from data extraction. AI tools play a crucial role not only to extract and structure data but also to fill in missing data points, thus supporting researchers in experiments and characterization.

A unified, FAIR-compliant database (see Figure 2) will significantly improve data accessibility and reuse by making standardized MOF data readily available to researchers worldwide. This enhanced accessibility facilitates collaboration and innovation, ultimately accelerating the development of MOF materials for CCU applications.

5. Conclusion

MOFs hold immense potential for advancing carbon capture and utilization technologies due to their customizable structures and exceptional properties. However, the field faces significant challenges that hinder rapid progress. The lack of standardized data and measurement protocols leads to inconsistencies in reporting key performance metrics, making it difficult to compare and optimize different MOF materials effectively. Additionally, the complexity and interconnected nature of CCU research require researchers to process vast amounts of data from diverse sources, which can be time-consuming and inefficient.

Developing a unified, open-access database dedicated to MOF materials for CCU applications would address these challenges by providing standardized and easily accessible data. Such a database would facilitate direct comparison of MOF materials, enable the identification of promising candidates, and accelerate innovation in the field. Recent advancements in NLP and AI, particularly LLMs, have made it possible to automate the extraction of

valuable data from the extensive body of scientific literature. These tools significantly reduce the difficulty and time required for data extraction, making the creation and maintenance of a comprehensive database feasible now.

By harnessing AI and NLP technologies to populate the database, researchers can more efficiently navigate the complex research landscape. Automated data extraction allows for the continuous updating of the database with the latest findings, ensuring that researchers have access to the most current information. This development not only streamlines the research process but also empowers scientists to manage and share their data effectively, adhering to FAIR principles.

In summary, addressing the challenges in MOF research for CCU applications requires a concerted effort to standardize methodologies and promote data sharing. The convergence of advanced AI tools and the establishment of a unified database make it possible to overcome existing barriers. These innovations facilitate coordinated and collaborative efforts within the CCU community, which could greatly accelerate the development of new MOF materials for carbon capture, utilization, and storage, ultimately contributing significantly to global efforts in combating climate change.

Acknowledgements

The authors acknowledged the Young Investigator Network (YIN) of the Karlsruhe Institute of Technology (KIT) for providing additional financial support through the YIN Award 2022. M.T. acknowledges support from the DACStorE project, funded by the Initiative and Networking Fund of the Helmholtz Association (grant agreement number KA2-HSC-12) and support by a Feodor Lynen Experienced Researcher Fellowship from the Alexander von Humboldt Foundation. The authors acknowledge the use of OpenAI's ChatGPT for language editing and refining parts of the text of this article.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Claudia Bizzarri: conceptualization (equal); funding acquisition (equal); and writing—original draft (equal). **Manuel Tsotsalas:** conceptualization (equal); funding acquisition (equal); and writing—original draft (equal).

Keywords

automated data extractions, carbon capture and conversions, FAIR (FAIR—Findable, Accessible, Interoperable, and Reusable) data principles, metal-organic frameworks, natural language processing tools, photocatalytic CO₂ reduction

Received: October 2, 2024

Revised: December 11, 2024

Published online:

[1] A. Leiserowitz, J. Carman, N. Buttermore, S. Rosenthal, K. Mulcahy, International Public Support for Climate Action, **2021**.

- [2] a) H. McLaughlin, A. A. Littlefield, M. Menefee, A. Kinzer, T. Hull, B. K. Sovacool, M. D. Bazilian, J. Kim, S. Griffiths, *Renewable Sustainable Energy Rev.* **2023**, *177*, 113215; b) H.-J. Ho, A. Iizuka, E. Shibata, *Ind. Eng. Chem. Res.* **2019**, *58*, 8941; c) F. Nocito, A. Dibenedetto, *Curr. Opin. Green Sustainable Chem.* **2020**, *21*, 34; d) M. Yousaf, M. Zaman, A. Mahmood, M. Imran, A. Elkamel, M. Rizwan, T. Wilberforce, F. Riaz, *Energy Sci. Eng.* **2022**, *10*, 4890.
- [3] N. S. Lewis, *Science* **2016**, *351*, aad1920.
- [4] J. Schneider, H. Jia, J. T. Muckerman, E. Fujita, *Chem. Soc. Rev.* **2012**, *41*, 2036.
- [5] <https://new.nsf.gov/funding/data-management-plan#nsfs-data-sharing-policy-1c8>.
- [6] DFG Guidelines on the Handling of Research Data, <https://www.dfg.de/resource/blob/172098/4ababf7a149da4247d018931587d76d6/guidelines-research-data-data.pdf>.
- [7] <https://www.rsc.org/journals-books-databases/author-and-reviewer-hub/authors-information/prepare-and-format/data-sharing/>.
- [8] a) M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. Bonino da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, *Sci. Data* **2016**, *3*, 160018; b) C.-L. Lin, P.-C. Huang, S. Graessle, C. Grathwol, P. Tremouilhac, S. Vanderheiden, P. Hodapp, S. Herres-Pawlis, A. Hoffmann, F. Fink, G. Manolikakes, T. Opatz, A. Link, M. M. B. Marques, L. J. Daumann, M. Tsotsalas, F. Biedermann, H. Mutlu, E. Täuscher, F. Bach, T. Drees, S. Neumann, N. Jung, S. Bräse, *ChemRxiv*, **2023**, <https://doi.org/10.26434/chemrxiv-2023-2dd4c>.
- [9] H. Mashhadimoslem, M. A. Abdol, P. Karimi, K. Zanganeh, A. Shafeen, A. Elkamel, M. Kamkar, *ACS Nano* **2024**, *18*, 23842.
- [10] a) M. Scheffler, M. Aeschlimann, M. Albrecht, T. Bereau, H. J. Bungartz, C. Felser, M. Greiner, A. Groß, C. T. Koch, K. Kremer, W. E. Nagel, M. Scheidgen, C. Wöll, C. Draxl, *Nature* **2022**, *604*, 635; b) C. Draxl, M. Scheffler, *J. Phys. Mater.* **2019**, *2*, 036001.
- [11] Materials Genome Initiative, Materials Genome Initiative, <https://www.mgi.gov/>, (accessed: September 2024).
- [12] a) R. Freund, S. Canossa, S. M. Cohen, W. Yan, H. Deng, V. Guillermin, M. Eddaoudi, D. G. Madden, D. Fairen-Jimenez, H. Lyu, L. K. Macreadie, Z. Ji, Y. Zhang, B. Wang, F. Haase, C. Wöll, O. Zaremba, J. Andreo, S. Wuttke, C. S. Diercks, *Angew. Chem. Int. Ed.* **2021**, *60*, 23946; b) H. Furukawa, K. E. Cordova, M. O'Keeffe, O. M. Yaghi, *Science* **2013**, *341*, 1230444; c) S. Kitagawa, R. Kitaura, S.-I. Noro, *Angew. Chem. Int. Ed.* **2004**, *43*, 2334.
- [13] B. Dziejarski, J. Serafin, K. Andersson, R. Krzyżyńska, *Mater. Today Sustainable* **2023**, *24*, 100483.
- [14] a) A. Samanta, A. Zhao, G. K. H. Shimizu, P. Sarkar, R. Gupta Ind, *Eng. Chem. Res.* **2012**, *51*, 1438; b) X. Shi, H. Xiao, H. Azarabadi, J. Song, X. Wu, X. Chen, K. S. Lackner, *Angew. Chem. Int. Ed.* **2020**, *59*, 6984.
- [15] a) K. Sumida, D. L. Rogow, J. A. Mason, T. M. McDonald, E. D. Bloch, Z. R. Herm, T.-H. Bae, J. R. Long, *Chem. Rev.* **2012**, *112*, 724; b) S. Bose, D. Sengupta, T. M. Rayder, X. Wang, K. O. Kirlikovali, A. K. Sekizkardes, T. Islamoglu, O. K. Farha, *Adv. Funct. Mater.* **2023**, 2307478.
- [16] a) R. Li, W. Zhang, K. Zhou, *Adv. Mater.* **2018**, *30*, 1705512; b) Z. Wang, H. Fei, Y.-N. Wu, *ChemSusChem* **2024**, *17*, e202400504.
- [17] R. Mao, H. Chai, S. Xie, J. Fan, X. Yu, *Sci. Adv.* **2024**, *10*, ead14449.
- [18] X. Feng, Y. Pi, Y. Song, C. Brzezinski, Z. Xu, Z. Li, W. Lin, *J. Am. Chem. Soc.* **2020**, *142*, 690.
- [19] a) R. Huang, Y. Peng, C. Wang, Z. Shi, W. Lin, *Eur. J. Inorg. Chem.* **2016**, 2016, 4358; b) W. Liang, M. Z. M. Noor, W. Lewis, A. Macmillan, X. Zhang, R. Tang, A. Kochubei, D. Wiley, B. S. Haynes, J. Huang, *ChemPhotoChem* **2023**, *7*, e202300031; c) H. Nagatomi, L. C. Gallington, S. Goswami, J. Duan, K. W. Chapman, N. Yanai, N. Kimizuka, O. K. Farha, J. T. Hupp, *ACS Omega* **2020**, *5*, 30299; d) X. Deng, J. Albero, L. Xu, H. García, Z. Li, *Inorg. Chem.* **2018**, *57*, 8276; e) P. M. Stanley, C. Thomas, E. Thyrhaug, A. Urstöger, M. Schuster, J. Hauer, B. Rieger, J. Warnan, R. A. Fischer, *ACS Catal.* **2021**, *11*, 871.
- [20] a) K. Zhang, S. Goswami, H. Noh, Z. Lu, T. Sheridan, J. Duan, W. Dong, J. T. Hupp, *J. Photochem. Photobiol.* **2022**, *10*, 100111; b) A. Aziz, A. R. Ruiz-Salvador, N. C. Hernández, S. Calero, S. Hamad, R. Grau-Crespo, *J. Mater. Chem. A* **2017**, *5*, 11894.
- [21] R. Harii, S. Dehghanpour, *Appl. Organomet. Chem.* **2021**, *35*, e6422.
- [22] S. Xie, C. Deng, Q. Huang, C. Zhang, C. Chen, J. Zhao, H. Sheng, *Angew. Chem. Int. Ed.* **2023**, *62*, e202216717.
- [23] a) N. Sadeghi, S. Sharifnia, M. Sheikh Arabi, *J. CO₂ Util.* **2016**, *16*, 450; b) E.-X. Chen, M. Qiu, Y.-F. Zhang, Y.-S. Zhu, L.-Y. Liu, Y.-Y. Sun, X. Bu, J. Zhang, Q. Lin, *Adv. Mater.* **2018**, *30*, 1704388.
- [24] a) Y. Chen, D. Wang, X. Deng, Z. Li, *Catal. Sci. Technol.* **2017**, *7*, 4893; b) M. Ding, R. W. Flaig, H.-L. Jiang, O. M. Yaghi, *Chem. Soc. Rev.* **2019**, *48*, 2783; c) M. Khan, Z. Akmal, M. Tayyab, S. Mansoor, A. Zeb, Z. Ye, J. Zhang, S. Wu, L. Wang, *Carbon Capture Sci. Technol.* **2024**, *11*, 100130; d) D. Li, M. Kassymova, X. Cai, S.-Q. Zang, H.-L. Jiang, *Coord. Chem. Rev.* **2020**, *412*, 213262; e) R. Mori, *Mater. Renewable Sustainable Energy* **2024**, *13*, 109; f) Q. Su, Q. Guo, H. Wang, M. Liu, C. Zuo, *Carbon Resour. Convers.* **2024**, *7*, 100049; g) G. Ding, C. Li, L. Chen, G. Liao, *Energy Environm. Sci.* **2024**, *17*, 5311; h) R. Mori, *Mater. Renewable Sustainable Energy* **2024**, *13*, 109; i) K. Sun, Y. Qian, H.-L. Jiang, *Angew. Chem. Int. Ed.* **2023**, *62*, e202217565.
- [25] a) D. Ziegenbalg, A. Pannwitz, S. Rau, B. Dietzek-Ivansic, C. Streb, *Angew. Chem. Int. Ed.* **2022**, *61*, e202114106; b) S. Fang, M. Rahaman, J. Bharti, E. Reisner, M. Robert, G. A. Ozin, Y. Hang-Hu, *Nat. Rev. Methods Primers* **2023**, *3*, 61; c) M. Bonchio, J. Bonin, O. Ishitani, T.-B. Lu, T. Morikawa, A. J. Morris, E. Reisner, D. Sarkar, F. Toma, M. Robert, *Nat. Catal.* **2023**, *6*, 657.
- [26] a) P. A. Saenz Cavazos, E. Hunter-Sellers, P. Iacomini, S. R. McIntyre, D. Danaci, D. R. Williams, *Front. Energy Res.* **2023**, *11*, 1167043; b) J. D. Evans, V. Bon, I. Senkova, S. Kaskel, *Langmuir* **2021**, *37*, 4222; c) A. J. Howarth, A. W. Peters, N. A. Vermeulen, T. C. Wang, J. T. Hupp, O. K. Farha, *Chem. Mater.* **2017**, *29*, 26.
- [27] C. Charalambous, E. Moubarak, J. Schilling, E. Sanchez Fernandez, J.-Y. Wang, L. Herraiz, F. Mcilwaine, S. B. Peh, M. Garvin, K. M. Jablonka, S. M. Moosavi, J. Van Herck, A. Y. Ozturk, A. Pourghaderi, A.-Y. Song, G. Mouchaham, C. Serre, J. A. Reimer, A. Bardow, B. Smit, S. Garcia, *Nature* **2024**, 632, 89.
- [28] a) M. Karimi, M. Shirzad, J. A. C. Silva, A. E. Rodrigues, *Environm. Chem. Lett.* **2023**, *21*, 2041; b) R. L. Siegelman, E. J. Kim, J. R. Long, *Nat. Mater.* **2021**, *20*, 1060; c) L. Jiang, W. Liu, R. Q. Wang, A. Gonzalez-Diaz, M. F. Rojas-Michaga, S. Michailos, M. Pourkashanian, X. J. Zhang, C. Font-Palma, *Prog. Energy Combust. Sci.* **2023**, *95*, 101069; d) Z. Zhang, Z. P. Cano, D. Luo, H. Dou, A. Yu, Z. Chen, *J. Mater. Chem. A* **2019**, *7*, 20985; e) D. M. Glenna, A. Jana, Q. Xu, Y. Wang, Y. Meng, Y. Yang, M. Neupane, L. Wang, H. Zhao, J. Qian, S. W. Snyder, *J. Phys.*

- Chem. Lett.* **2023**, 14, 10693; f) M.-Y. Low, L. V. Barton, R. Pini, C. Petit, *Chem. Engin. Res. Des.* **2023**, 189, 745; g) A. Kumar, D. G. Madden, M. Lusi, K.-J. Chen, E. A. Daniels, T. Curtin, J. J. Perry, M. J. Zaworotko, *Angew. Chem. Int. Ed.* **2015**, 23, 14372; h) X. Shi, H. Xiao, H. Azarabadi, J. Song, X. Wu, X. Chen, K. S. Lackner, *Angew. Chem. Int. Ed.* **2020**, 59, 6984; i) D. Fu, M. E. Davis, *Chem. Soc. Rev.* **2022**, 51, 9340.
- [29] L. Espinal, D. L. Poster, W. Wong-Ng, A. J. Allen, M. L. Green, *Environ. Sci. Technol.* **2013**, 47, 11960.
- [30] A. H. Farmahini, S. Krishnamurthy, D. Friedrich, S. Brandani, L. Sarkisov, *Chem. Rev.* **2021**, 121, 10666.
- [31] J. Young, F. McIlwaine, B. Smit, S. Garcia, M. van der Spek, *Chem. Eng. J.* **2023**, 456, 141035.
- [32] a) G. H. Han, J. Bang, G. Park, S. Choe, Y. J. Jang, H. W. Jang, S. Y. Kim, S. H. Ahn, *Small* **2023**, 19, e2205765; b) J. Albero, Y. Peng, H. García, *ACS Catal.* **2020**, 10, 5734; c) C. Li, J. Wang, L. Tong, Y. Wang, P. Zhang, M. Zhu, H. Dong, *Coord. Chem. Rev.* **2024**, 502, 215623; d) Y. Cui, A. Labidi, X. Liang, X. Huang, J. Wang, X. Li, Q. Dong, X. Zhang, S. I. Othman, A. A. Allam, D. W. Bahnemann, C. Wang, *ChemSusChem* **2024**, 17, e202400551.
- [33] a) H. Takeda, C. Cometto, O. Ishitani, M. Robert, *ACS Catal.* **2017**, 7, 70; b) C. Bizzarri, *Eur. J. Org. Chem.* **2022**, 2022, e202200185; c) M. Khalil, J. Gunazuardi, T. A. Ivandini, A. Umar, *Renew. Sustain. Energy Rev.* **2019**, 113, 109246.
- [34] A. Behera, A. K. Kar, R. Srivastava, *Mater. Horiz.* **2022**, 9, 607.
- [35] International Union of Pure and Applied Chemistry (IUPAC), *Compendium of Chemical Terminology*, 3.0.1 ed., **2019**, <https://doi.org/10.1351/goldbook.T06534>.
- [36] S. Kozuch, J. M. L. Martin, *ACS Catal.* **2012**, 2, 2787.
- [37] a) H. Kisch, D. Bahnemann, *J. Phys. Chem. Lett.* **2015**, 6, 1907; b) R. Ricka, M. Přibyl, K. Kočí, *Appl. Catal. A Gen.* **2023**, 658, 119166.
- [38] S. E. Braslavsky, A. M. Braun, A. E. Cassano, A. V. Emeline, M. I. Litter, L. Palmisano, V. N. Parmon, N. Serpone, *Pure Appl. Chem.* **2011**, 83, 931.
- [39] B. Reiß, E. Riedle, H.-A. Wagenknecht, *ChemPhotoChem* **2021**, 5, 1009.
- [40] U. Caudillo-Flores, M. J. Muñoz-Batista, M. Fernández-García, A. Kubacka, *Catal. Rev.* **2022**, 66, 531.
- [41] International Union of Pure and Applied Chemistry (IUPAC), *Compendium of Chemical Terminology*, 3.0.1 ed., **2019**, <https://doi.org/10.1351/goldbook.S05564>.
- [42] a) S. Navarron-Jaén, M. Virginie, J. Bonin, M. Robert, R. Wojcieszak, A. Y. Khodakov, *Nat. Rev. Chem.* **2021**, 5, 564; b) Z. Guo, G. Chen, C. Cometto, B. Ma, H. Zhao, T. Groizard, L. Chen, H. Fan, W.-L. Man, S.-M. Yiu, K.-C. Lau, T.-C. Lau, M. Robert, *Nat. Catal.* **2019**, 2, 801; c) L.-L. Gracia, P. Henkel, O. Fuhr, C. Bizzarri, *Beilstein J. Org. Chem.* **2023**, 19, 1766; d) F. Droghetti, A. Amati, F. Pascale, A. Crochet, M. Pastore, A. Ruggi, M. Natali, *ChemSusChem* **2024**, 17, e202300737.
- [43] a) D. Siderius, NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials, National Institute of Standards and Technology, <https://doi.org/10.18434/T43882>, (accessed: September 2024); b) X. Cai, F. Gharagheizi, L. W. Bingel, D. Shade, K. S. Walton, D. S. Sholl, *Ind. Eng. Chem. Res.* **2021**, 60, 639.
- [44] J. D. Evans, V. Bon, I. Senkovska, S. Kaskel, *Langmuir* **2021**, 37, 4222.
- [45] P. Tremouilhac, C.-L. Lin, P.-C. Huang, Y.-C. Huang, A. Nguyen, N. Jung, F. Bach, R. Ulrich, B. Neumair, A. Streit, S. Bräse, *Angew. Chem. Int. Ed.* **2020**, 59, 22771.
- [46] Y. G. Chung, E. Haldoupis, B. J. Bucior, M. Haranczyk, S. Lee, H. Zhang, K. D. Vogiatzis, M. Milisavljevic, S. Ling, J. S. Camp, B. Slater, J. I. Siepmann, D. S. Sholl, R. Q. Snurr, *J. Chem. Eng. Data* **2019**, 64, 5985.
- [47] N. S. Bobbitt, K. Shi, B. J. Bucior, H. Chen, N. Tracy-Amoroso, Z. Li, Y. Sun, J. H. Merlin, J. I. Siepmann, D. W. Siderius, R. Q. Snurr, *J. Chem. Eng. Data* **2023**, 68, 483.
- [48] a) P. Kalhor, N. Jung, S. Bräse, C. Wöll, M. Tsotsalas, P. Friederich, *Adv. Funct. Mater.* **2024**, 34, 2302630; b) M. Schilling-Wilhelmi, M. Ríos-García, S. Shabih, M. V. Gil, S. Miret, C. T. Koch, J. A. Márquez, K. M. Jablonka, *2024*, arXiv:2407.16867, <https://doi.org/10.48550/arXiv.2407.16867>; c) J. van Herck, M. V. Gil, K. M. Jablonka, A. Abrudan, A. Anker, M. Asgari, B. Blaiszik, A. Buffo, L. Choudhury, C. Corminboeuf, H. Daglar, A. M. Elahi, I. T. Foster, S. Garcia, M. Garvin, G. Godin, L. L. Good, J. Gu, N. X. Hu, X. Jin, T. Junkers, S. Keskin, T. P. J. Knowles, R. Laplaza, M. Lessona, S. Majumdar, H. Mashhadimoslem, R. D. McIntosh, S. M. Moosavi, B. Mourño, et al., *Chem. Sci.* **2025**, <https://doi.org/10.1039/d4sc04401k>; d) J. Zhao, S. Huang, J. M. Cole, *J. Chem. Inf. Model.* **2023**, 63, 1961; e) E. J. Beard, J. M. Cole, *Sci. Data* **2022**, 9, 329.
- [49] a) A. Nandy, C. Duan, H. J. Kulik, *J. Am. Chem. Soc.* **2021**, 143, 17535; b) P. Z. Moghadam, S. M. J. Rogge, A. Li, C.-M. Chow, J. Wierme, N. Moharrami, M. Aragonés-Anglada, G. Conduit, D. A. Gomez-Gualdrón, V. Van Speybroeck, D. Fairen-Jimenez, *Matter* **2019**, 1, 219; c) G. G. Terrones, S.-P. Huang, M. P. Rivera, S. Yue, A. Hernandez, H. J. Kulik, *J. Am. Chem. Soc.* **2024**, 146, 20333.
- [50] a) Z. Zheng, O. Zhang, C. Borgs, J. T. Chayes, O. M. Yaghi, *J. Am. Chem. Soc.* **2023**, 145, 18048; b) M. Jalali, M. Tsotsalas, C. Wöll, *Nanomaterials* **2022**, 12, 704; c) Y. Luo, S. Bag, O. Zaremba, A. Cierpka, J. Andreo, S. Wuttke, P. Friederich, M. Tsotsalas, *Angew. Chem. Int. Ed.* **2022**, 61, e202200242; d) Y. Luo, S. Bag, O. Zaremba, A. Cierpka, J. Andreo, S. Wuttke, P. Friederich, M. Tsotsalas, *Angew. Chem.* **2022**, 134, e202200242; e) M. Rahimi, S. M. Moosavi, B. Smit, T. A. Hatton, *Cell Rep. Phys. Sci.* **2021**, 2, 100396.
- [51] a) K. M. Jablonka, P. Schwaller, A. Ortega-Guerrero, B. Smit, *Nat. Mach. Intell.* **2024**, 6, 161; b) M. Caldas Ramos, C. J. Collison, A. D. White, *2024*, arXiv:2407.01603.



Claudia Bizzarri was appointed as an associate professor at the University of Rome Tor Vergata in September 2024. Previously, she worked as an independent group leader and lecturer at the Karlsruhe Institute of Technology, where she obtained the *Habilitation (venia legendi)* in organic chemistry. During this time, she developed earth-abundant-metal-based photoactive coordination complexes for photocatalytic CO₂ reduction, photochemistry, and bioimaging. Her current research continues to address sustainability challenges, particularly in the areas of solar energy utilization and waste valorization.



Manuel Tsotsalas is an experimental chemist specializing in materials science, with applications in health and sustainability. He earned his Ph.D. under Prof. Luisa De Cola at the University of Münster and conducted postdoctoral research with Prof. Susumu Kitagawa at Kyoto University. Currently, he is a tenured group leader and lecturer at the Karlsruhe Institute of Technology (KIT) and a visiting lecturer at Northwestern University, where he collaborates on combining experimental and computational workflows for material discovery. As cofounder and CTO of Aixelo Inc., he works to integrate AI-powered decision-making into research, promoting FAIR data practices in the field.