2022

Artificial intelligence-based material discovery for clean energy future

Reza Maleki

Mohsen Asadnia

Amir Razmjou
Edith Cowan University

10.1002/aisy.202200073
Maleki, R., Asadnia, M., & Razmjou, A. (2022). Artificial intelligence-based material discovery for clean energy future. Advanced Intelligent Systems, 4(10), Article 2200073. https://doi.org/10.1002/aisy.202200073

ADVANCED INTELLIGENT SYSTEMS
Open Access
www.advintellsyst.com

Artificial Intelligence-Based Material Discovery for Clean Energy Future

*Reza Maleki, Mohsen Asadnia, and Amir Razmjou**

Artificial intelligence (AI)-assisted materials design and discovery methods can come to the aid of global concerns for introducing new efficient materials in different applications. Also, a sustainable clean future requires a transition to a low-carbon economy that is material-intensive. AI-assisted methods advent as inexpensive and accelerated methods in the design of new materials for clean energies. Herein, the emerging research area of AI-assisted material discovery with a focus on developing clean energies is discussed. The applications, advantages, and challenges of using AI in material discovery are discussed and the future perspective of using AI in clean energy is studied. This perspective paves the way for a better understanding of the future of AI applications in clean energies.

1. Introduction on Materials Discovery for Clean Energy

Increasing restrictions on the use of fossil fuels to mitigate the environmental implications of air pollution and global warming have encouraged countries around the globe to find alternative clean energy resources.[1,2] **Figure 1**a shows the timeline of the clean energies progresses worldwide. Several new methods have been introduced for harnessing energy resources with less pollution and renewability. Efforts to reduce fossil fuel consumption have led to the development of new technologies such as an electric vehicle (EV) powered by lithium-ion batteries, thermoelectric materials, fuel cells, photovoltaics (PV), etc.[3] These technologies require a large number of materials and minerals. For example, a typical EV's battery has over 6,000 individual lithium-ion cells with a total weight of around 500 kg, which consists of about 11.5 kg lithium, 27 kg nickel, 20 kg of manganese, 13.5 kg cobalt, 91 kg copper, and 180 kg aluminum, steel, and plastic. Extraction of one ton of lithium carbonate equivalent (LCE) from ore (spodumene) produces at least 15.8 tons of CO_2, and for brine the value reduces to around 0.3 tons of CO_2 (33.9 kg CO_2 eq. per kWh for an NMC111 chemistry battery).[4–6] Water footprint for brine is around 470 tons of water per ton of lithium and the value for rock mining is around 170 tons. The development of clean energy technologies and processes requires the discovery of new materials for increasing process efficiency, reducing carbon, water, and land footprints, as well as minimizing capital expenditures (CAPEX) and operating expenses (OPEX).

Discovering the new materials using conventional approaches requires significant financial and time investments. Evaluating patents reveals that it takes approximately 1–2 two decades from the discovery of new material to its first commercial use.[7] The rapidly growing demand for global clean energy has imposed significant pressure on research institutes to accelerate the discovery of advanced materials that can be used in the swift implementation of clean energy processes.

2. Role of Artificial Intelligence

One of the best ways to speed up material discovery is through using artificial intelligence (AI). The application of AI methods in data analysis has increased thanks to the significant advancement in the development of supercomputers.[8] AI has been used in various fields such as energy, environment, material science, management, economy, etc.[7,9–11] It is one of the most important tools that has been recently used to study and accelerate the deployment of clean energy technologies/materials. AI-assisted methods are known as the fourth paradigm in the materials design and discovery.[12] Machine learning (ML) algorithms, as

R. Maleki
Department of Chemical Engineering
Shiraz University
Shiraz 71345, Iran

M. Asadnia
School of Engineering
Macquarie University
Sydney, NSW 2109, Australia

A. Razmjou
School of Engineering
Edith Cowan University
Joondalup, Perth, WA 6027, Australia
E-mail: amirr@unsw.edu.au

A. Razmjou
UNESCO Centre for Membrane Science and Technology
School of Chemical Engineering
University of New South Wales
Sydney, NSW 2052, Australia

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202200073.

DOI: 10.1002/aisy.202200073

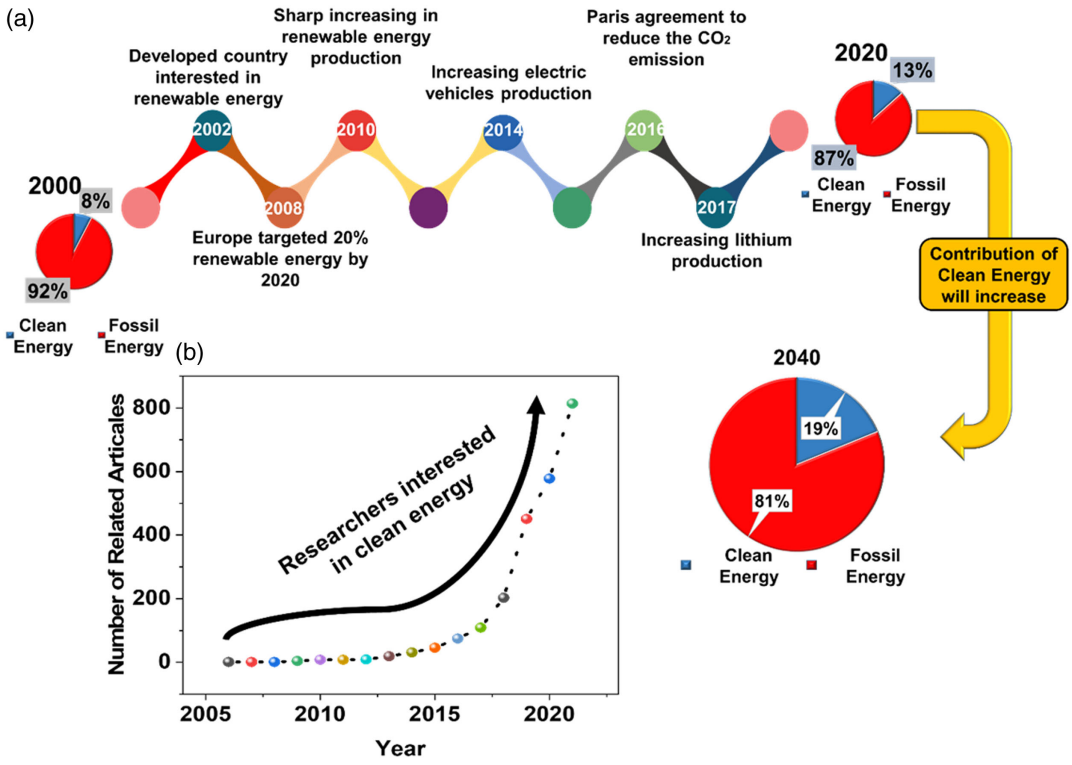

Figure 1. a) Timeline of the contribution of clean energy in total produced energy worldwide and b) the number of published articles related to AI-based materials discovery for clean energy.[22–26]

an important group of AI-assisted methods for the prediction of materials properties, materials design, and materials optimization, are categorized into the four main types: supervised learning, unsupervised learning, and reinforcement learning. The main difference between supervised and unsupervised learning is that in the first, the data are labeled and the input and output of each model are determined; however, in the latter, the data are unlabeled and we have tried to find and establish a relationship between input values. The ordinary least square regression (OLSR), support vector regression (SVR), Gaussian process regression (GPR), kernel ridge regression (KRR), and decision tree (DT) are many examples of supervised learning. On the other hand, clustering and dimensionality reduction are two common approaches to unsupervised learning.[13] Also, other methods such as artificial neural network (ANN) and deep learning are utilized frequently as AI-based methods for materials design.[14]

Compared with traditional lab-based trial-and-error material discovery, machine learning (ML), a branch of AI, has shown the ability to cut short the material development time. The conventional experimental measurements rely mainly on the intuition of researchers over a long period, which suffers greatly from human error. Material developments solely based on computational simulations are micro-/nanostructure dependent and can predict the individual building block properties of materials and are inaccurate postassembly, not to mention that the simulation requires expensive computing clusters/supercomputers. However, AI-assisted material discovery does not have the drawbacks of traditional methods and can facilitate material development in an unprecedented inexpensive/rapid way.

Among different AI-based methods, their advantages and limitations should be noticed carefully. For instance, take machine learning and deep learning as two important AI-assisted methods. Deep learning (DL) or deep materials informatics is a modern version of neural network algorithms. Thanks to the recent development and availability of enough data and computational infrastructures, deep learning finds promising real-world applications (e.g., self-driving cars). Deep learning possesses some principal advantages. First, deep learning is known as a feature-engineering independent technique due to its potential power in extracting related features. This point not only helps to save time which is used for the feature engineering process in traditional machine learning, but also provides a good opportunity to find new features. Second, since in deep learning a massive volume of datasets are trained, deep learning predictive models are more reliable and accurate rather than traditional machine learning models. Third, again because of using big data in deep learning, their prediction is more accelerated than machine learning predictive models. However, providing the required big data for deep learning is one of the serious challenges. Also, big data require more time to be trained. In addition, due to the complexity of deep learning models, they work like a black box in which there is any evidence about what exactly happened in them.[15] In brief, some advantages and limitations of AI methods are shown in **Table 1**.

Table 1. Advantages and limitations of most commonly used AI approaches.

Name of the AI-based methods	Advantages (In the Materials Science)	Limitations (In the Materials Science)
Machine learning	Can train on small datasets	Less accuracy
	Required less time to train	The necessity to try different features and algorithms to achieve the best prediction
	Can be done on CPU	
		Require highly accurate preprocessing
Deep learning	Able to extract related features automatically	Require big data
		Computationally intensive
		Time-consuming training procedure
	Preprocessing is not necessary	Require big computing hardware (GPU)
	Building more accurate and reliable models	Highly complex and not easily interpretable
	Fast prediction	
Artificial neural network (ANN)	Good prediction capacity	Hard to interpret due to its blackbox nature
	Applicable for nonmathematical models	
	Easy to use	Less accuracy
	Able to model difficult functions	Extreme complexity due to network structure
	Able to capture the interaction between predictors	Require accurate training
		Require long processing time
	Possibility to apply multiple different training algorithms	Sustainable to overfitting (due to complexity)

As shown in Figure 1b, the number of AI-based researches on clean energy is growing exponentially. There are many comprehensive case studies[16–19] and review papers that have investigated the progress of using AI-assisted methods for designing renewable materials (e.g., batteries).[20,21] For instance, Lombardo et al.[21] focused on the advantages, and challenges of using AI/ML in batteries. Also, based on electrochemical results, ML would be able to identify reaction mechanisms that lead to enhanced experimental knowledge, which is related to batteries. However, using AI/ML in batteries confronted many challenges, for example, unavailability of descriptors, which can be generalized for different problems, imbalanced data and overfitting issues, and lack of strong collaboration between experimental and computational experts in the battery field. In another review, Chen et al.[20] investigated the applications of the ML as an emerged tool for use in rechargeable batteries. In this review, they explained how ML can utilize collected information in experimental and computational datasets. In material science, a powerful AI tool can widely analyze and improve the properties of different materials by analyzing previous material characteristics obtained from various experiments and computational chemistry.

AI builds a model using data from a "limited" number of experimental or computational studies on material properties. The model, then, will help to predict the properties of an "unlimited" number of materials. Thus, using AI predictions, the need to perform experimental or computational studies will be reduced sustainability, which in turn will accelerate the material discovery. There are many examples of using AI in different groups of materials in the Materials Science field such as renewable energy materials, silicon-based materials, PV materials, lithium ions electrodes, etc.

Ho Gu et al.[26] reviewed the capabilities of AI in data analysis of renewable energy materials and its role in reducing their production cost by a factor of 10 for running experimental tests and simulations. In a review article, Pollice et al.[27] proposed AI as a useful tool to predict the properties of silicon-based materials (such as silicon–germanium alloys and perovskite/c-silicon composites) to be used in PV. Wang et al.[28] tried to overcome the two important challenges of reducing in power generation costs of PV materials and increasing their efficiency. One of the important groups of PVs is chalcogenide semiconductors with photoelectric properties. In this study, with the tackle to the wide search space of chalcogenide semiconductor structures and via using 1520 quaternary chalcogenide, eight quaternary chalcogenide materials were screened via unsupervised learning. The selected chalcogenide materials were simulated through quantum mechanics (ab initio calculations) and the results showed that eight PV candidates have the desired photoelectric properties. These results suggested the potential power of ML models in the discovery of efficient quaternary chalcogenide semiconductors in a short period of time in comparison with the other discovery methods that will take long, near 12.1 years.

The widespread use of wind and solar energies has been limited due to some of their intrinsic issues mainly discontinuous energy production, which requires energy storage systems such as batteries to provide continuous energy delivery from the batteries to the consumer devices. The evergrowing demand for EVs (sales of EVs in the global market have increased by an average of 60% annually since 2014[29]) also requires a large deployment of the batteries as they have high efficiency and energy storage capacity.[30–32] The high level of demand has made the role of the batteries pivotal in the transition to zero-carbon civilization. Gao et al.[33] reviewed the use of AI as an effective method for the

preparation of new electrodes for lithium-ion batteries such as hollow porous carbon materials, nickel manganese, and cobalt oxide. In addition, they emphasized the significant role of electrode designing and its undeniable effect on the batteries' performance.

One of the main challenges in discovering new materials for batteries is a large number of proposed materials for electrodes and electrolytes. Screening these materials is expensive, time-consuming, and requires expensive infrastructure, which makes the evaluation of the new materials for use in lithium batteries impractical. Accordingly, researchers are looking for fast ways to discover or optimize materials for energy storage applications.[34–36] The use of AI makes it possible to consider simultaneously a large volume of information related to material properties and characterizations. AI also provides a chance to screen effective parameters for determining new materials.

Zhang et al.[37] suggested AI as a way to discover new sulfur host materials for lithium–sulfur batteries. Their technique shortened the materials screening time by ≈8 years. In their study, using 2DMatPedia database, 14 new materials have been discovered that have the potential to be utilized in lithium–sulfur batteries. In their AI-aided material screening, the adsorption energies of the lithium polysulfides on the materials were considered as the main criterion. Their AI-based model has been developed using a small dataset including 65 samples. This dataset contained the adsorption energies of lithium polysulfides on materials such as PdN_2, TaS_2, PtN_2, $TaSe_2$, $AgCl_2$, $NbSe_2$, $TaTe_2$, AgF_2, NiN_2, AuS_2, TmI_2, $NbTe_2$, $NiBi_2$, and $AuBr_2$. Using the developed model, the adsorption energies on the new materials were predicted.

Lv et al.[38] reviewed the applications of AI in the electrolyte (such as propylene carbonate and ethylene carbonate) and electrode (such as $ACoO_2$, $ANiO_2$, $ATiO_2$, etc.) materials developments for lithium-ion batteries. They studied the use of clustering, regression, and classification methods in discovering the new materials. By reviewing the applications of AI in battery materials, they have concluded that AI is a promising tool to speed up the discovery of materials for use in lithium-ion batteries, especially for electrolytes and electrodes. Similarly, Liu et al.[39] reviewed the role of AI in developing lithium-ion batteries and highlighted that the AI models can predict the performance of different electrolytes and electrodes materials and suggest their optimum conditions. Xu et al.[40] emphasized the microstructures characteristics of the lithium-ion battery as an important factor that governs batteries' efficacy. They showed that AI can contribute to predicting the performance of batteries with certain microstructure designs, so it can help select the best design at low cost with minimal effort. To pave the way for greater use of Zn-ion batteries, using AI, Zhou et al.[41] proposed new cathode materials such as $SnPO_4$, $MnPO_4$, $AgPS_3$, and CoP_2O_7 for Zn-ion batteries.

3. Limitations of Artificial Intelligence

One of the limitations of AI for studying battery exploration materials and other technologies is the low size of datasets. AI models require large amounts of data and extensive data banks to work properly.[42] With the launch of the Materials Genome Initiative (MGI) in 2011 and the coming of the "big data" era,[43] several publicly available datasets were initiated, which provide researchers with access to the properties of known and hybrid materials. A few examples of data banks are Harvard Clean Energy Project (HCEP),[44] Cambridge Structural Databases,[45] inorganic crystal structure database (ICSD),[46] Materials Project (MP),[47] Materials Data Facility,[48] and open quantum materials database (OQMD).[49] A comprehensive list of many examples of materials datasets which are used in the renewable energy applications is summarized in **Table 2**.

Apart from using available materials database for materials design and discovery, there are other kinds of data capturing methods such as natural language processing (NLP) and text mining, which are used in different areas of materials discovery, for example, design of battery materials.[58] Text mining techniques can extract unstructured data hidden in published documents (e.g., papers, patents, datasheets, and reports) and turn them into structured materials data.[77] This extraction procedure can be done via different chemistry-aware NLP toolkits such as OSCAR4,[78] tmChem,[79] and ChemDataExtractor.[80] Huang et al.[81] via using ChemDataExtractor on 229 061 academic documents extracted a battery database that contained 292 313 data. This huge amount of data is public and can be used for prediction properties in the field of batteries. Another case study, which was done by Kononova et al.,[82] extracted and collected a dataset of 19 488 solid-state synthesis recipes (codified recipes) from scientific documents via using ChemDataExtractor. This dataset contains a wide range of information (e.g., kinds of operation and related experimental conditions, initializing compounds, target structure) that can be used widely in the field of inorganic materials.

Since batteries and other clean energy systems are relatively new technologies, there are no related data banks nor enough experimental data to train and test the AI models. Fortunately, the advent of supercomputers has provided a great opportunity for advanced computational chemistry and quantum mechanics such as molecular dynamics (MD), coarse-grained MD, and density functional theory (DFT) to predict precisely material properties at the molecular/atom level.[83] As presented in **Figure 2**, computational chemistry and quantum mechanics can generate synthetic data that is complementary to the available experimental data that AI needs for testing and training its models. Jha et al.[84] compared the mean absolute error (MAE) of two different ML models for the prediction of materials formation energy. These deep learning models (OQMD-Scratch and transfer learning) are trained via four different datasets that three of them, OQMD, Materials Project, and JARVIS, were properties that were computed via DFT-based methods and one experimental dataset (formation energies of the 1643 samples). They compared the MAE of the deep transfer learning and MAE of OQMD-Scratch with the experimental dataset, which was 0.07 and ≈0.15 eV atom^{-1}, respectively. Their results suggested the advantage of using deep transfer learning for predicting properties, which are closer to their values that are obtained via experimental investigation.

These deep level computations can create suitable databases for AI-based discovery of new materials and develop technologies for the production of clean energy.[85] AI itself can also assist computational chemistry and quantum mechanics by predicting

Table 2. Some of the materials databases with detailed information.

Name of database	Website	Free/ purchasable	Related description		Reference[s]
ZINC	https://zinc.docking.org/	Purchasable	230 million compounds: 3D formats, genes, and chemical compounds		[50]
Materials Project (MP)	https://materialsproject.org/	Free	Inorganic compounds	Elastic tensors	[47]
			Band structures	Piezoelectric	
			Molecules	tensors	
			Nanoporous materials	Intercalation electrodes	
				Conversion electrodes	
AFLOWLIB	http://aflowlib.org/	Free	Elastic properties	Band structures	[51]
			Thermal properties	Binary, ternary, quaternary systems	
			Superalloys		
Novel materials discovery (NOMAD)	https://nomad-lab.eu/about/scope	Free	Computational materials: input and output files from more than 100 million calculations		[52]
Computational materials repository (CMR)	https://cmr.fysik.dtu.dk/	Free	Monolayer transition metal dichalcogenides and oxides	$ABSe_3$ materials	[53]
			van der Waals heterostructures	2D materials	
			Organometal halide perovskites	A_2BCX_4 materials	
				ABS_3 materials	
Inorganic crystal structure database (ICSD)	https://icsd.products.fiz-karlsruhe.de/	Free	Inorganic crystal structures		[54]
Harvard clean energy project (HCEP)	https://dash.harvard.edu/handle/1/8364968	Purchasable	Power conversion efficiency		[44]
			Energy of the highest occupied molecular orbital		
			Energy of the lowest occupied molecular orbital		
Harvard organic photovoltaic dataset (HOPV15)	https://dash.harvard.edu/handle/1/29408375	–	Experimental PV data		[55]
Pauling file database	https://paulingfile.com/	Purchasable	Electronic and electrical properties	Phase transitions	[56]
			Ferroelectric properties	Superconductor properties	
			Magnetic properties	Thermal and thermodynamic properties	
			Mechanical properties	Optical properties	
Citrination	https://citrination.com/	Free	Materials Data Platform		[57]
Quantum Machines 9 (QM9 database)/ GDB-13,	http://quantum-machine.org/datasets/	N/A	Computed geometric		[58,59]
			Energetic, electronic, and thermodynamic properties		
			134 k stable small organic molecules		
			1 billion stable and synthetically accessible organic molecules		
Crystallographic open database (COD)	http://www.crystallography.net/cod/	Free	Crystal structures		[56]
			Organic, inorganic, metal-organics compounds		
Joint automated repository for various integrated simulations (JARVIS)	https://jarvis.nist.gov/	Free	40 000 materials		[60]
			1 million calculated properties		
			500 materials		
			110 force-fields		
			25 ML models for material-property predictions		
Polymer Genome (PG)	https://www.polymergenome.org/	Free	Polymers		[61-63]
ASM Alloy phase diagram database	https://www.asminternational.org/	Purchasable	40 300 binary and ternary alloy phase diagrams		[64]
ChemSpider	http://www.chemspider.com/	Free	Chemical structure database		[65]

Table 2. Continued.

Name of database	Website	Free/ purchasable	Related description	Reference[s]
High Performance Alloy D atabase (HPAD)	https://cindasdata.com/products/hpad	Purchasable	46 900 curves on 137 alloys	[66]
SpringerMaterials	https://materials.springer.com/	Purchasable	Inorganic solid phases Metal foam Polymer thermodynamics	[67]
Open quantum materials database (OQMD)	https://oqmd.org/	Free	Database of DFT calculated thermodynamic and structural properties of 815 654 materials	[49,68]
Materials Cloud	www.materialscloud.org	Free	22'145'443 crystal structures	[69]
Materials Platform for Data Science (MPDS)	https://www.mpds.io/	Purchasable	Optical properties; Magnetic properties; Phase transitions; Electronic and electrical properties; Superconductivity; Thermal and thermodynamic properties; Mechanical properties	[70]
Material Properties Open Database (MPOD)	http://mpod.cimav.edu.mx/	Free	Physical properties of crystalline materials	[70]
American Mineralogist Crystal Structure Database (AMCSD)	http://rruff.geo.arizona.edu/AMS/amcsd.php	Free	Mineral structures	[70]
Bilbao server	http://www.cryst.ehu.es	Free	Structural and magnetic properties	[71]
IZA Zeolite database	http://www.iza-structure.org/databases/	Free	Structural information of Zeolite Framework types	[72]
NanoHUB	https://nanohub.org/	Free	Nanoelectronics; Photonics; Materials	
MAGNDATA	http://webbdcrista1.ehu.es/magndata/	Free	Magnetic structures	[73,74]
Predicted Crystallography Open Database (PCOD)	http://www.crystallography.net/pcod/	Free	Inorganic compounds	[75]
Theoretical Crystallography Open Database (TCOD)	http://www.crystallography.net/tcod/	Free	Theoretically calculated or refined crystal structures of: organic, inorganic, metal-organic compounds and minerals, excluding biopolymers	[75]
Cambridge crystallographic data center (CCDC)	https://www.ccdc.cam.ac./uk	Purchasable	Organic structures Metal-organic structures	[64]
Biological Macromolecule Crystallization Database (BMCD)	http://bmcd.ibbr.umd.edu/	Free	Molecule Crystal Crystallization data for macromolecules	[76]

Figure 2. Computational chemistry and quantum mechanics can provide missing data or complete the available datasets for training and testing AI models for material discovery.

the result of the simulation before it has been performed. Using the previous result of similar simulations, AI can predict the result of the simulations to select the best simulation conditions, which reduce the DFT and MD time and errors.[86] Lv et al.[38] obtained a sufficient dataset of materials properties using DFT and MD. They have identified these databases as an effective factor in developing the use of AI in the discovery of new materials for developing clean energy applications. With AI/DFT,

Park et al.[87] introduced P′3-type $K_{0.3}Mn_{0.9}Cu_{0.1}O_2$ (KMCO) as an electrode for potassium-ion batteries (KIBs).

Thermoelectric materials are another group of materials with wide potential application in the clean energy applications. In thermoelectric materials, via applying the electrical current flow, temperature difference and heat creates. This mechanism makes them suitable for recovery of the waste heat as well as refrigeration application. AI-assisted materials design and discovery methods are applied on thermoelectric materials, too.[88–90] Since the dopants have a significant role in improving thermoelectric performance, Na et al.[90] suggested a specific neural network (DopNet) for the prediction of the thermoelectric properties in the doped materials. In this study, they measured the effects of different dopants on the five thermoelectric properties (e.g., Seebeck coefficient, thermal conductivity, electrical conductivity, power factor, and figure of merit). For instance, the MAE of 0.06 for the figure of merit prediction indicated that DopNet could be a successful model in the prediction of efficient thermoelectric materials.

4. Challenges and Future Outlook

We are living in an era where global warming and climate change has already impacted the Earth. Urgent short- and long-term actions must be taken into consideration to mitigate the environmental implications. Abandoning fuel-based economy and transiting to a mineral-based one and clean energy future require widespread commercial implementation of green technologies and processes that can harness energy from nonfossil fuel-based energy resources. This urgent transition cannot be achieved without the rapid discovery of new materials that serves the technologies/process. AI can catalyze the material development and optimization which facilitates the transition to industry 4.0.[3,91,92] It seems that the main obstacle to AI-assisted material discovery is the lack of datasets and accurate models that can predict and suggest new materials.[93–95] Advanced computational chemistry and quantum mechanics can assist in addressing the bottleneck by creating synthetic data. We believe research needs to be directed in the following areas. 1) Developing new AI models that can predict a broad range of organic and inorganic materials as well as hybrid/composite materials. 2) Creating publicly available material data sets that can be used as the main source of model training and testing. 3) There is an urgent need in creating a bridge language that facilitates communication between the computer and material scientists. 4) AI-assisted material discovery requires a well-defined model input and output parameters which require a scientific consensus among the researchers to report their results under a universal framework systematically and consistently. 5) MD and DFT simulations are currently based on predicting the properties of materials at the atom/molecular level which might be enough to suggest building blocks/starting materials. However, they become inaccurate in predicting the bulk of materials or after assembling the building blocks in the desired morphology. There is a need in developing MD-/DFT-like approaches that can predict the building block properties, postassembly. 6) To achieve sustainability for technology materials production, AI material discovery models should consider the environmental impact as an important contributory factor. This will provide a chance for the new material to be linked with life cycle assessments (LCAs) performance data, and good quality social and governance data within the supply chains.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

artificial intelligence, energy, machine learning, material discovery, sustainability

Received: March 14, 2022
Revised: June 3, 2022
Published online: August 4, 2022

[1] D. J. Davidson, *Nat. Energy* **2019**, *4*, 254.
[2] R. York, *Nat. Clim. Change* **2012**, *2*, 441.
[3] D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwaraknath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, A. Aspuru-Guzik, *Nat. Rev. Mater.* **2018**, *3*, 5.
[4] Y. Bai, T. Zhang, Y. Zhai, Y. Jia, K. Ren, J. Hong, *J. Environ. Manage.* **2022**, *312*, 114949.
[5] S. Jiang, L. Zhang, F. Li, H. Hua, X. Liu, Z. Yuan, H. Wu, *J. Environ. Manage.* **2020**, *262*, 110253.
[6] D. A. Notter, M. Gauch, R. Widmer, P. Wager, A. Stamp, R. Zah, H.-J. Althaus, *Contribution of Li-Ion Batteries to the Environmental Impact of Electric Vehicles*, ACS Publications, Washington, DC, **2010**.
[7] M. Fetanat, M. Keshtiara, Z.-X. Low, R. Keyikoglu, A. Khataee, Y. Orooji, V. Chen, G. Leslie, A. Razmjou, *Ind. Eng. Chem. Res.* **2021**, *60*, 5236.
[8] R. Pollice, G. Dos Passos Gomes, M. Aldeghi, R. J. Hickman, M. Krenn, C. Lavigne, M. Lindner-D'Addario, A. K. Nigam, C. T. Ser, Z. Yao, A. Aspuru-Guzik, *Acc. Chem. Res* **2021**, *54*, 849.
[9] Z. Li, S. Wang, H. Xin, *Nat. Catal.* **2018**, *1*, 641.
[10] M. Chai, S. Moradi, E. Erfani, M. Asadnia, V. Chen, A. Razmjou, *Chem. Mater.* **2021**, *33*, 8666.
[11] M. Fetanat, M. Keshtiara, R. Keyikoglu, A. Khataee, R. Daiyan, A. Razmjou, *Sep. Purif. Technol.* **2021**, *270*, 118383.
[12] R. Jose, S. Ramakrishna, *Appl. Mater. Today* **2018**, *10*, 127.
[13] Y. Juan, Y. Dai, Y. Yang, J. Zhang, *J. Mater. Sci. Technol.* **2021**, *79*, 178.
[14] J. Schmidt, M. R. G. Marques, S. Botti, M. A. L. Marques, *npj Comput. Mater.* **2019**, *5*, 1.
[15] A. Agrawal, A. Choudhary, *MRS Commun.* **2019**, *9*, 779.
[16] N. Yao, X. Chen, X. Shen, R. Zhang, Z.-H. Fu, X.-X. Ma, X.-Q. Zhang, B.-Q. Li, Q. Zhang, *Angew. Chem., Int. Ed.* **2021**, *60*, 21473.
[17] G. Wang, T. Fearn, T. Wang, K.-L. Choy, *ACS Cent. Sci.* **2021**, *7*, 1551.
[18] G. Chaney, A. Ibrahim, F. Ersan, D. Çakır, C. Ataca, *ACS Appl. Mater. Interfaces* **2021**, *13*, 36388.
[19] M. Kunitski, N. Eicke, P. Huber, J. Köhler, S. Zeller, J. Voigtsberger, N. Schlott, K. Henrichs, H. Sann, F. Trinter, L. P. H. Schmidt, A. Kalinin, M. S. Schöffler, T. Jahnke, M. Lein, R. Dörner, *Nat. Commun.* **2019**, *10*, 1.
[20] X. Chen, X. Liu, X. Shen, Q. Zhang, *Angew. Chem., Int. Ed.* **2021**, *60*, 24354.

[21] T. Lombardo, M. Duquesnoy, H. El-Bouysidy, F. Årén, A. Gallo-Bueno, P. B. Jørgensen, A. Bhowmik, A. Demortière, E. Ayerbe, F. Alcaide, *Chem. Rev.* **2021**, *122*, 10899.
[22] *EIA Petroleum and Natural Gas Predicted to Remain Top of Sources of World Energy Consumption Through 2040*. https://www.michiganoilandgas.org/petroleum_and_natural_gas_predicted_to_remain.
[23] *BP Statistical Review of World Energy*, **2020**.
[24] Energy Renewable energy in Europe. https://ec.europa.eu/info/news/focus-renewable-energy-europe-2020-mar-18_en.
[25] N. Vukajlović, D. Milićević, B. Dumnić, B. Popadić, *J. Energy Storage* **2020**, *31*, 101603.
[26] G. H. Gu, J. Noh, I. Kim, Y. Jung, *J. Mater. Chem. A* **2019**, *7*, 17096.
[27] R. Pollice, G. Dos Passos Gomes, M. Aldeghi, R. J. Hickman, M. Krenn, C. Lavigne, M. Lindner-D'Addario, A. Nigam, C. T. Ser, Z. Yao, A. Aspuru-Guzik, *Acc. Chem. Res.* **2021**, *54*, 849.
[28] Q.-J. Hong, J. Schroers, D. Hofmann, S. Curtarolo, M. Asta, A. Van De Walle, *npj Comput. Mater.* **2021**, *7*, 1.
[29] W. Wen, S. Yang, P. Zhou, S. Gao, *Renewable Sustainable Energy Rev.* **2021**, *144*, 111024.
[30] J. Dunn, M. Slattery, A. Kendall, H. Ambrose, S. Shen, *Environ. Sci. Technol.* **2021**, *55*, 5189.
[31] J. C. Barbosa, R. Gonçalves, C. M. Costa, S. Lanceros-Mendez, *Energies* **2021**, *14*, 3145.
[32] X. Chen, T. Hou, K. A. Persson, Q. Zhang, *Mater. Today* **2019**, *22*, 142.
[33] X. Gao, X. Liu, R. He, M. Wang, W. Xie, N. P. Brandon, B. Wu, H. Ling, S. Yang, *Energy Storage Mater.* **2021**, *36*, 435.
[34] E. O. Pyzer-Knapp, K. Li, A. Aspuru-Guzik, *Adv. Funct. Mater.* **2015**, *25*, 6495.
[35] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, *npj Comput. Mater.* **2017**, *3*, 1.
[36] P. Srivastava, R. Abbassi, A. K. Yadav, V. Garaniya, M. Asadnia, *Chemosphere* **2020**, *254*, 126926.
[37] H. Zhang, Z. Wang, J. Cai, S. Wu, J. Li, *ACS Appl. Mater. Interfaces* **2021**, *13*, 53388.
[38] C. Lv, X. Zhou, L. Zhong, C. Yan, M. Srinivasan, Z. W. Seh, C. Liu, H. Pan, S. Li, Y. Wen, Q. Yan, *Adv. Mater.* **2021**, *34*, 2101474.
[39] Y. Liu, Q. Zhou, G. Cui, *Small Methods* **2021**, *5*, 2100442.
[40] H. Xu, J. Zhu, D. P. Finegan, H. Zhao, X. Lu, W. Li, N. Hoffman, A. Bertei, P. Shearing, M. Z. Bazant, *Adv. Mater.* **2021**, *11*, 2003908.
[41] L. Zhou, A. M. Yao, Y. Wu, Z. Hu, Y. Huang, Z. Hong, *Adv. Theor. Simul.* **2021**, *4*, 2100196.
[42] Y. Duan, J. S. Edwards, Y. K. Dwivedi, *Int. J. Inf. Manage.* **2019**, *48*, 63.
[43] Y. Liu, T. Zhao, W. Ju, S. Shi, *J. Materiomics* **2017**, *3*, 159.
[44] J. Hachmann, R. Olivares-Amaya, S. Atahan-Evrenk, C. Amador-Bedolla, R. S. Sánchez-Carrera, A. Gold-Parker, L. Vogt, A. M. Brockway, A. Aspuru-Guzik, *J. Phys. Chem. Lett.* **2011**, *2*, 2241.
[45] I. A. Baburin, V. A. Blatov, L. Carlucci, G. Ciani, D. M. Proserpio, *Cryst. Growth Des.* **2008**, *8*, 519.
[46] L. Grünewald, M. Langer, S. Meyer, D. Nerz, J. Hänisch, B. Holzapfel, D. Gerthsen, *Supercond. Sci. Technol.* **2021**, *34*, 035005.
[47] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
[48] B. Blaiszik, K. Chard, J. Pruyne, R. Ananthakrishnan, S. Tuecke, I. Foster, *JOM* **2016**, *68*, 2045.
[49] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, C. Wolverton, *npj Comput. Mater.* **2015**, *1*, 1.
[50] T. Sterling, J. J. Irwin, *J. Chem. Inf. Comput. Sci.* **2015**, *55*, 2324.
[51] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L.W. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, O. Levy, *Comput. Mater. Sci.* **2012**, *58*, 227.
[52] *The Novel Materials Discovery (NOMAD) Laboratory - NOMAD Lab*, https://nomad-lab.eu/about/scope.
[53] *CMR — Computational Materials Repository*, https://cmr.fysik.dtu.dk/, n.d.
[54] *How to use | ICSD*" https://icsd.products.fiz-karlsruhe.de/en/howuse/how-use, n.d.
[55] S. A. Lopez, E. O. Pyzer-Knapp, G. N. Simm, T. Lutzow, K. Li, L. R. Seress, J. Hachmann, A. Aspuru-Guzik, *Sci. data* **2016**, *3*, 1.
[56] Pauling File, https://paulingfile.com/, n.d.
[57] Citrination, https://citrination.com/users/sign_in, n.d.
[58] L. Ruddigkeit, R. Van Deursen, L. C. Blum, J.-L. Reymond, *J. Chem. Inf. Comput. Sci.* **2012**, *52*, 2864.
[59] R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. Von Lilienfeld, *Sci. Data* **2014**, *1*, 1.
[60] K. Choudhary, K. F. Garrity, A. C. Reid, B. DeCost, A. J. Biacchi, A. R. Hight Walker, Z. Trautt, J. Hattrick-Simpers, A. G. Kusne, A. Centrone, *npj Comput. Mater.* **2020**, *6*, 1.
[61] H. Doan Tran, C. Kim, L. Chen, A. Chandrasekaran, R. Batra, S. Venkatram, D. Kamal, J. P. Lightstone, R. Gurnani, P. Shetty, M. Ramprasad, J. Laws, M. Shelton, R. Ramprasad, *J. Appl. Phys.* **2020**, *128*, 171104.
[62] C. Kim, A. Chandrasekaran, T. D. Huan, D. Das, R. Ramprasad, *J. Phys. Chem. C* **2018**, *122*, 17575.
[63] T. D. Huan, A. Mannodi-Kanakkithodi, C. Kim, V. Sharma, G. Pilania, R. Ramprasad, *Sci. Data* **2016**, *3*, 1.
[64] *Home - The Cambridge Crystallographic Data Centre (CCDC)*, https://www.ccdc.cam.ac.uk/, n.d.
[65] J. Zhang, D. Gao, Q. Li, Y. Zhao, L. Li, H. Lin, Q. Bi, Y. Zhao, *Sci. Total Environ.* **2020**, *704*, 135931.
[66] *High Performance Alloys Database | CINDAS LLC*, https://cindasdata.com/products/hpad, n.d.
[67] *SpringerMaterials – Properties of Materials*, https://materials.springer.com/, n.d.
[68] B. R. Jany, A. Janas, W. Piskorz, K. Szajna, A. Kryshtal, G. Cempura, P. Indyka, A. Kruk, A. Czyrska-Filemonowicz, F. Krok, *Nanoscale* **2020**, *12*, 9067.
[69] L. Talirz, S. Kumbhar, E. Passaro, A. V. Yakutovich, V. Granata, F. Gargiulo, M. Borelli, M. Uhrin, S. P. Huber, S. Zoupanos, C. S. Adorf, C. W. Andersen, O. Schütt, C. A. Pignedoli, D. Passerone, J. Vandevondele, T. C. Schulthess, B. Smit, G. Pizzi, N. Marzari, *Sci. Data* **2020**, *7*, 1.
[70] *MPDS Materials Platform*, https://mpds.io/#start, n.d.
[71] M. I. Aroyo, J. M. Perez-Mato, C. Capillas, E. Kroumova, S. Ivantchev, G. Madariaga, A. Kirov, H. Wondratschek, *Z. Kristallogr.-Cryst. Mater.* **2006**, *221*, 15.
[72] *Database of Zeolite Structures*, http://www.iza-structure.org/databases/, n.d.
[73] S. V. Gallego, J. M. Perez-Mato, L. Elcoro, E. S. Tasci, R. M. Hanson, M. I. Aroyo, G. Madariaga, *J. Appl. Crystallogr.* **2016**, *49*, 1941.
[74] S. V. Gallego, J. M. Perez-Mato, L. Elcoro, E. S. Tasci, R. M. Hanson, K. Momma, M. I. Aroyo, G. Madariaga, *J. Appl. Crystallogr.* **2016**, *49*, 1750.
[75] B. Coimbatore Narayanan, J. Westbrook, S. Ghosh, A. I. Petrov, B. Sweeney, C. L. Zirbel, N. B. Leontis, H. M. Berman, *Nucleic Acids Res.* **2014**, *42*, D114.
[76] J. E. Lee, M. L. Fusco, D. M. Abelson, A. J. Hessell, D. R. Burton, E. O. Saphire, *Indian J. Pharm. Sci.* **2009**, *65*, 1162.
[77] E. O. Pyzer-Knapp, J. W. Pitera, P. W. J. Staar, S. Takeda, T. Laino, D. P. Sanders, J. Sexton, J. R. Smith, A. Curioni, *npj Comput. Mater.* **2022**, *8*, 1.
[78] D. M. Jessop, S. E. Adams, E. L. Willighagen, L. Hawizy, P. Murray-Rust, *J. Cheminform.* **2011**, *3*, 1.
[79] R. Leaman, C.-H. Wei, Z. Lu, *J. Cheminform.* **2015**, *7*, 1.
[80] M. C. Swain, J. M. Cole, *J. Chem. Inf. Comput. Sci.* **2016**, *56*, 1894.

[81] S. Huang, J. M. Cole, *Sci. Data* **2020**, *7*, 1.
[82] M. Torres-Beltrán, A. K. Hawley, D. Capelle, E. Zaikova, D. A. Walsh, A. Mueller, M. Scofield, C. Payne, L. Pakhomova, S. Kheirandish, J. Finke, M. Bhatia, O. Shevchuk, E. A. Gies, D. Fairley, C. Michiels, C. A. Suttle, F. Whitney, S. A. Crowe, P. D. Tortell, S. J. Hallam, *Sci. Data* **2019**, *6*, 1.
[83] A. Razmjou, M. Asadnia, O. Ghaebi, H.-C. Yang, M. Ebrahimi Warkiani, J. Hou, V. Chen, *ACS Appl. Mater. Interfaces* **2017**, *9*, 38076.
[84] D. Jha, K. Choudhary, F. Tavazza, W.-K. Liao, A. Choudhary, C. Campbell, A. Agrawal, *Nat. Commun.* **2019**, *10*, 1.
[85] J. Li, K. Lim, H. Yang, Z. Ren, S. Raghavan, P.-Y. Chen, T. Buonassisi, X. Wang, *Matter* **2020**, *3*, 393.
[86] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, K.-R. Müller, *Chem. Rev.* **2021**, *121*, 10142.
[87] S. Park, S. Park, Y. Park, M. H. Alfaruqi, J.-Y. Hwang, J. Kim, *Energy Environ. Sci.* **2021**, *14*, 5864.
[88] T. Wang, C. Zhang, H. Snoussi, G. Zhang, *Adv. Funct. Mater.* **2020**, *30*, 1906041.
[89] P. Gorai, V. Stevanović, E. S. Toberer, *Nat. Rev. Mater.* **2017**, *2*, 1.
[90] G. S. Na, S. Jang, H. Chang, *npj Comput. Mater.* **2021**, *7*, 1.
[91] R. Batra, L. Song, R. Ramprasad, *Nat. Rev. Mater.* **2021**, *6*, 655.
[92] R. W. Epps, A. A. Volk, K. G. Reyes, M. Abolhasani, *Chem. Sci.* **2021**, *12*, 6025.
[93] J. Cowls, A. Tsamados, M. Taddeo, L. Floridi, *Nat. Mach. Intell.* **2021**, *3*, 111.
[94] A. Karargyris, S. Kashyap, I. Lourentzou, J. T. Wu, A. Sharma, M. Tong, S. Abedin, D. Beymer, V. Mukherjee, E. A. Krupinski, M. Moradi, *Sci. Data* **2021**, *8*, 1.
[95] A. C. Vaucher, P. Schwaller, J. Geluykens, V. H. Nair, A. Iuliano, T. Laino, *Nat. Commun.* **2021**, *12*, 1.