CATALYSIS

Catalyst: New Materials Discovery: Machine-Enhanced Human Creativity

Geoffrey Ozin[1,*] and Todd Siler[1]

Geoffrey Ozin is a Distinguished University Professor at the University of Toronto and Government of Canada Research Chair in Materials Chemistry and Nanochemistry. He is renowned for his work in defining, enabling, and popularizing a chemical approach to nanomaterials for innovative nanotechnology in advanced materials and biomedical science.

Dr. Todd Siler is a multimedia artist trained in psychology and art at the Massachusetts Institute of Technology. His published works include *Breaking the Mind Barrier* and *Think like a Genius*. For the past four decades, Siler has exhibited his artworks internationally with representation by Ronald Feldman Fine Arts in New York City.

With the exception of the laws of nature—or "natural laws," such as Newton's laws of motion, Boyle's law, Charles's law, Gay-Lussac's law, and thermodynamic laws—nature makes no patented claims of final limits on anything the human mind can conceive, comprehend, and construct for any and all purposes. We should not be surprised to learn that the most sophisticated sentient machines of our own making might know this truth while we humans might never know (Figure 1).

Given the constant feats of our most advanced artificial intelligence (AI) technology in a range of fields, we might want to reconsider how machines can prove algorithmically superior to humans at screening and optimizing the properties of massive libraries of known materials and reducing the time to market of a product or process. They could even help us invent completely new markets and economies that we have overlooked or missed—or simply never considered—because we have willfully blindsided ourselves to the unforeseen possibilities of human potential. As "lifelong learners," our best attempt at a fully realized AI might one day enjoy an infinite lifespan that no mortal will experience.

This article, seen through the eyes of a long-time practicing chemist and an artist, asks whether machine learning will ever be creative enough to match the innate ingenuity of humans at discovering and synthesizing an entirely new class of materials (Figure 2).

It takes an overarching knowledge of materials chemistry augmented by solid-state physics, tremendous insight, innate creativity, instinctive synthetic skills, and a little luck to discover a class of materials never before known to humans. By "discover," we do not mean facilitate an accelerated incremental technical improvement to the properties of a known class of materials by using computational, compositional, and structural combinatorial chemistry. Instead, we refer to the invention (through rationally designed chemical synthesis rather than trial and error that might lead to lucky discoveries) of a novel material structure and composition that departs from the prevailing paradigm—celebrating the inimitability of human ingenuity, delighting the senses, unveiling relationships between structure and properties, heralding novel functionality, and enabling new technologies.

Along this vein, Figure 3 represents the five tenets of solid-state materials chemistry. They have provided generations of chemists with a tried-and-true invaluable guide to the intelligently planned synthesis of purposeful materials for specific applications.

It is important to point out that this graphical representation, itself composed of words alone and loosely configured as a piece of "concrete poetry," is limited both in its visual vocabulary and in the way it engages the imagination. Unlike many thinking machines, human beings are able to welcome a wide range of ambiguity, abstraction, multiple interpretations, and related phenomena, all of which are central to the more emotionally charged, evocative imagery of metaphorical information that neuroscientists are discovering as the key to creative learning, elaborative encoding, long-term memory, and metacognition. These interrelated areas are central to the neuroscience of creativity and neuro-aesthetics, as well as to art-based methods and tools, which have proven highly beneficial to science, technology, engineering, and mathematics (STEM) education.

To replicate an already abstruse concept such as human creativity, advanced AI-enabled learning machines must be coded to engage in the open-minded blue-sky, divergent, and lateral thinking practices that are the hallmarks of invention and innovation. They have to be tasked to envision myriad ways of representing and communicating the interconnections between these five tenets: synthesis, structure, property, function, and utility. Each new representation offers the possibility of expanding our views on these tenets and the material possibilities that can arise out of their combination.

Programming machines to optimize their creative potential will inevitably work to advance the frontiers of materials chemistry, spurring new and productive cross-disciplinary collaborations, which will bring to bear essential scientific breakthroughs and technological innovations. Importantly, it is also very likely that we as humans stand

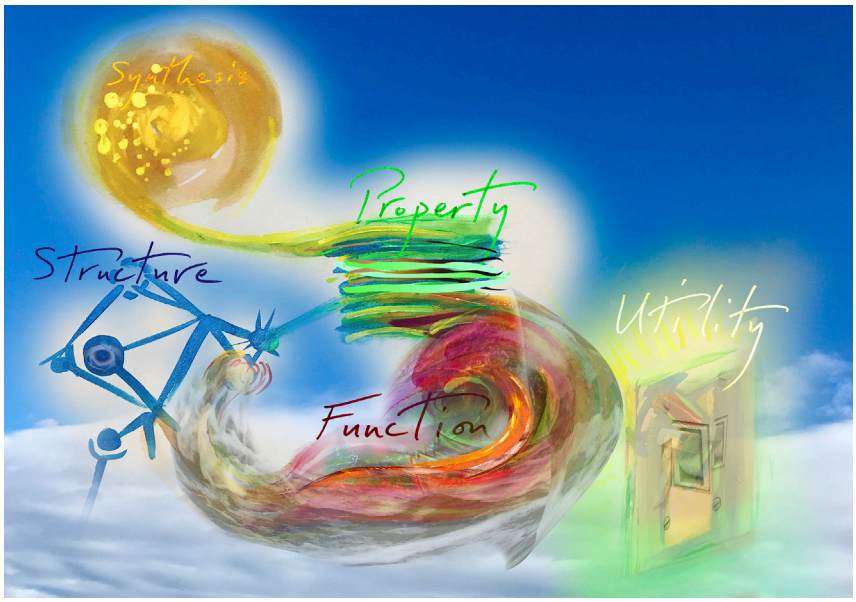

Figure 1. Metaphorming: Connecting and Transforming All Information in Personally Meaningful, Purposeful, Productive, and Useful Ways

The five tenets of solid-state materials chemistry: synthesis, structure, property, function, and utility. ArtScience graphic illustration courtesy of G.O. and T.S. (http://www.artnanoinnovations.com). ("ArtScience" represents a new domain of communication that fully integrates the methods and tools of creative inquiry and discovery used in the arts and sciences. The ArtScience process aims to connect and transform potentially all fields of human knowledge to catalyze and cultivate innovative thinking while improving human communication by deepening understanding.)

to gain insight into our own creative processes by interfacing with these machines. Ultimately, these learning machines could spark a new type of innovative thinking in humans, prompting us to understand the real and virtual or imagined connectivity of all information. If used properly, advanced AI-enabled learning machines will benefit humanity.

Primarily, without this art and science of applied imagination and knowledge, there would be no mathematics or physics and, in turn, no chemistry or biology!

The point is that materials discovery is a very human endeavor. It requires a grand synthetic and analytical chemistry skill set, innate chemical intuition, and the ability to recognize and exploit serendipity in research. Acquired through practice, this is a talent that, over time, automation and computers will have to match. In the end, machines might do more than miraculously demonstrate the occasional "eureka" moment. The current literature on materials science frequently boasts studies using machine learning to accelerate the discovery of all possible compositions of multi-metallic glasses, intermetallic compounds, and multi-element crystals. Materials innovation and computer epiphany might be on the horizon after all. The point is that we have a vast amount of data relating to materials created by the ingenuity of materials chemists. The amount is so vast that we cannot profit from it without assistance from learning machines.

More dramatically and profoundly, these algorithms could recognize and consistently apply some key mechanisms of human creativity, such as those illuminated by the two distinguished creativity researchers Robert and Michele Root-Bernstein, co-authors of *Sparks of Genius: The 13 Thinking Tools of the World's Most Creative People* (Houghton and Mifflin, 1999).

The first report of combinatorial inorganic materials chemistry occurred in the mid-1950s and was aimed at accelerating materials discovery. This breakthrough lay more or less dormant until the mid-1990s, when the field of combinatorics underwent a renaissance wherein practitioners extolled excitement, expectations, and promises of a technological revolution driven by materials discovery.

It turns out that creating vast libraries of inorganic materials by top-down and bottom-up methods, as well as rapidly testing them for specific property-function relationships, has yet to lead to the discovery of significant numbers of techno-economically relevant materials. That said, we might be missing or misrepresenting something fundamental to the discovery process in our approach to tapping the secret to materials innovation.

But, we must never say never. As a matter of course, we will inevitably be amazed and impressed if modern computational methods succeed in out-innovating years of synthetic combinatorial chemistry.

Other ArtScience explorers think otherwise. They propose conducting more basic thought experiments that challenge our present thinking on the limits of machine learning—specifically, testing our machines' limitations in rapidly fabricating, analyzing, and predicting the properties of massive libraries of known classes of materials.

AI ultimately has to learn from prior work based on established chemical principles and/or the published literature and extrapolate forward. Crystallographers were ahead of the game by developing properly curated databases

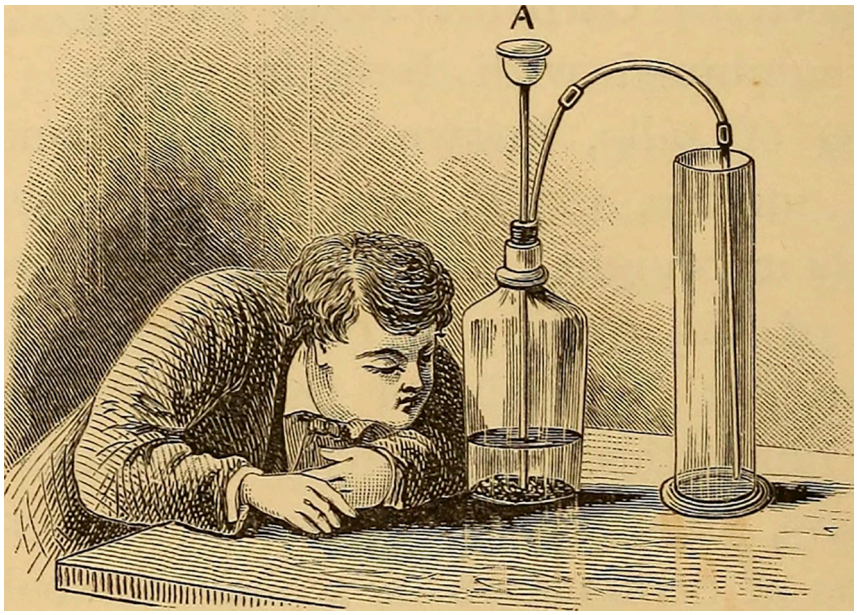

Figure 2. Woodcut Print Depicting a 19th Century Chemist Looking at, and Thinking about, His Possible New Discovery

This woodcut conveys the crucial role of human ingenuity in unearthing something new, a contrast to modern invention through automation, AI, machine learning, and big data and echoing the thesis of this article. Interestingly, the man in this woodcut is actually preparing CO_2. In every generation, some of the leaders will question whether modern techniques should be used instead of the tried and tested techniques passed on by previous generations. If new discovery-enabling algorithms had been available to previous generations, they would have been used to augment the innate ingenuity of humans for the benefit of humanity. Illustration courtesy of Wikimedia Commons.

in a common format; however, for materials properties, there are few reliable datasets, and none are comprehensive. One of the big challenges now is making all materials data accessible; otherwise, the revolution of materials discovery might never happen.

Clearly, machines can rapidly tweak elemental compositions to optimize the properties of materials we know, but they have yet to prove capable of discovering classes of materials we do not know. However, we do not know what we do not know. Is there a way out of this self-built intellectual prison? Certainly not without continually testing the various limits of our tacit and explicit knowledge. Considered together with the existing sum of human knowledge, that essential test will be ongoing and might remain incomplete by nature.

In the interest of testing this hypothesis, why not task our most visionary sentient machines with this single challenge: show and describe new materials, or classes of materials, that do not exist today. What new applications and utilities do you envision capitalizing on?

In the case of known-unknown materials (which we know should exist but have not yet been synthesized), 12 examples of which are offered later, discovery will present a challenge to which computational intelligence could very well rise. However, the discovery of unknown-unknown materials (which have been neither imagined nor synthesized) will require futuristic computational creativity techniques such as inverse design. By computer-aided inverse design, we mean choosing a specific property and function and screening for its structure instead of selecting a structure to determine its properties and function. An initial selection of a massive library of classes of materials that could fulfil these requirements is subject to a high-throughput theoretical search, which identifies promising candidate materials for high-throughput synthesis and characterization protocols. These steps provide a focused group of materials, further narrowed to the material of choice by circular synthesis, structure-property characterization, and theoretical procedures.

Developments in computational power and infrastructure, together with the recent emergence of machine learning, are already turning rapid screening of entire classes of molecules into real-world applications, especially in medicine through the discovery of new and efficacious pharmaceuticals. The rapid screening of materials from first principles could very well emerge after further improvements in computer performance and intelligence, especially if a creative synthetic materials chemist is directly involved in the development of the algorithm. Human versus AI is not the question. Humans and AI will always beat AI or humans on their own. So let the creative materials chemist become even more inventive and creative by using AI as a tool!

Therein lies the value proposition for exploring the art of science and the science in art. Yet, even if the most sophisticated set of computational tools could predict a brand-new class of materials, the question would remain as to how to synthesize it. Moreover, if it proves to be technologically significant, how then do we scale the synthesis to industrially relevant proportions?

Ultimately, the most accurate theoretical predictions will still rely on creatively orchestrated synthetic chemistry tactics to make them a reality. However, once human inventiveness unveils a new class of materials, machines can accelerate the time to market.

Figure 3. The Five Tenets of Solid-State Materials Chemistry

(1) Chemical synthesis creates new materials, (2) X-ray diffraction provides structural information, (3) electronic band theory elucidates properties, (4) defect type and population generate function, and (5) the relationships between structure, property, and function enable utility.

The aforementioned "synthesis gap" represents the bottleneck that currently exists once a target material has been identified in a machine-learning search. In principle, one can surmount this bottleneck by tasking an AI system, by using natural language-processing methodologies, to automatically mine information from vast libraries of publications that document all possible synthesis parameters and preparative strategies for producing particular classes of materials with a predefined size, shape, surface, defect type and population, degree of perfection, and state of aggregation. In addition to this list of synthesis requirements, it is important to include product yield, purity, reproducibility, and scale. Such a system could decide on an optimum recipe for making the target material and even the energy requirements and cost. The recent literature includes signs that this gigantic task might just happen.

In this context, an exquisite example of human ingenuity at the pinnacle of innovative synthetic chemistry is the recent discovery of a new material dubbed borophane (BH). This material is the hydrogenated form of borophene (B), a new member of the emerging class of two-dimensional materials that are inspired by the graphene archetype and that include the subsequently discovered transition-metal dichalcogenides, boron nitride, phosphorene, and silicene.

BH was discovered through an incredibly creative synthetic strategy.[1] First, the choice of precursor to BH was the layered material, magnesium diboride, MgB_2 (Figure 4). This material, known for its low-temperature superconductivity, is composed of planar B_2^{2-} layers and has a structure built from fused, six-atom boron rings interspersed with charge-balancing Mg^{2+} cations. The synthetic challenge was to extract the B sheets intact and then cap and stabilize them with charge-balancing hydrogen atoms while preventing adventitious hydrolysis or oxidation.

To accomplish the aforementioned clean proton-exchange reaction (i.e., $MgB_2 + 2H^+ \rightarrow Mg^{2+} + 2BH$) while avoiding unwanted side reactions, the authors mixed an ion-exchange resin with a non-aqueous methanol or acetonitrile suspension of MgB_2 under an inert N_2 atmosphere at room temperature and under ambient pressure.[1] The yellow filtrate was dried and subsequently shown to yield single-sheet BH with an impressive yield of 42.3%. Figure 4 shows the synthetic protocol for preparing the BH sheets alongside the structure of the sheets, which, according to extensive microscopy, diffraction, and spectroscopy analyses, are most likely composed of a double hydrogen-bridged, hexagonal boron network.

BH is a powerful demonstration of synthetic materials chemistry unveiling a new class of materials by intelligent design rather than a trial-and-error approach. This is a wonderful case study where human inventiveness and resourcefulness currently outstrip programmed robots and machines. Despite the prevailing outlook on the future of materials discovery through computational approaches, synthetically driven materials breakthroughs continue to surprise, delight, and still forge the path to the future.

Could a robot or machine have dreamt up the incredibly creative synthetic strategy reported for BH? Perhaps a machine-learning tool could if the computer scientist were to code the machine with an algorithm that applied the above-mentioned creative search-and-decision-making paradigm to make BH.

With a long history of miraculous breakthroughs in the art and science of materials discovery through chemistry, we

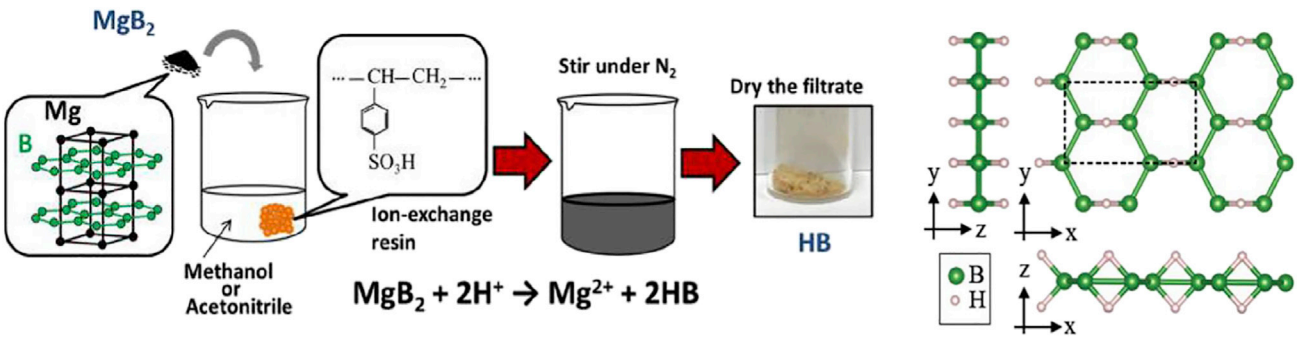

Figure 4. The Synthesis and Structure of Single-Sheet Borophane
Adapted with permission from Nishino et al.[1] Copyright 2017 American Chemical Society.

view a future of brute-force computational materials screening with guarded optimism. It is an idea that people like to toss around, yet the field currently has very little to show for itself in creating and accelerating new materials-based technologies from the typical 10 years of incubation to as little as 1–2 years.

In part, this stems from a general belief that in the future, computers and big data will play a role in all aspects of society, including science and technology. The current thinking on the future of materials discovery is detracting and distracting from important synthetic discoveries made by talented chemists, such as the discovery of BH as highlighted in this article, which need a public space to be recognized, openly discussed, and widely celebrated.

With the BH breakthrough now accomplished, creative chemistry aided by robots and machines can serve to accelerate the long road from discovery to market by expanding upon its structure-composition-property-function relationships and ultimately leading to applications and commercialization.

Recall the quote from Philip Ball, a past editor at *Nature*: "A breakthrough is a discovery pregnant with promise and then the hard graft begins." This certainly reflects the joy experienced by the synthetic chemist who discovers a brand new class of materials, at which point the tough work of materials optimization starts.

More to the point, a true "breakthrough" challenges all previous limitations of imagination, intuition, data, and knowledge. It counters and disrupts our present understanding of things with the evidence-based reality of a scientific, medical, or technological innovation.

In effect, a breakthrough embodies the essence of creativity. Specifically, a human (and perhaps soon also a machine) experiences and is able to express an unconditional response to familiar and unfamiliar things, stimuli, data, knowledge, events, and so forth to create something new, purposeful, productive, and useful on the basis of experience.

Machines, needless to say, can aid in this time-intensive process by cutting down on the effort required for humans to achieve a desired figure of merit for a particular application of a new material. But, none of it would be possible without a timely moment of human ingenuity and an astute recognition of serendipity.

The only concern we have with this development is whether the materials chemists of the future will be exposed to the same depth of experience as the current leading materials chemists in the field because they have grown up being too reliant on machines. As a result, they will lack or not have fully developed the stroke of genius that the current generation has.

Testing this thesis remains the challenge and boldest task of both human and machine intelligences combined: finding complementary ways of adventuring into the infinite frontier of innovation.

Holy Grails of Materials Chemistry: Challenges for AI, Machine Learning, Computer Intelligence, and Automation

There is an impressive history of materials discovery through chemistry, which has enabled the development of many notable transformative technologies in information technology, the generation and storage of clean energy, light-emitting diodes and lasers, sensors and detectors, transportation, agriculture, medicine, and health.

In the never-ending pursuit of "holy grail" materials that have profoundly influenced technologies, economies, and societies, those listed below continue to stretch the boundaries and test the pinnacles of "innovation wisdom" to reveal our individual and collaborative creativity and ingenuity. It will be most beneficial to see whether automation and machine learning can provide a helping hand and mind.

12 of these "holy grail" materials are as follows:

1. Room-temperature superconductor
2. N_2 fixation electrocatalyst
3. Material harder than diamond
4. Monodisperse nanocrystals
5. k < 1.5 interlayer dielectric
6. Practical artificial photosynthesis catalyst
7. Stable Si anode
8. Multicolor Si light-emitting diode
9. Perfect anti-counterfeit material
10. Single-step urea synthesis catalyst
11. Selective methane functionalization catalyst
12. Cold and hot fusion catalyst

Computer Creativity

If genius is a superhuman attribute enabling creativity and innovation in many different fields—an unbounded polymath personifying the universal mind, ceaseless curiosity, and knowledge of everything knowable—can we expect a computer algorithm, even if written by a genius programmer, to match or surpass human genius, which is a very human trait?

Human creativity is just the product of a neural network with connection weights on the order of about the square root of Avogadro's number. That is many thousands of times bigger that the neural networks currently trained. The architecture and learning procedures are selected by evolution, and the weights are trained for around a billion seconds with helpful inputs from other neural networks. There is no reason to believe that a machine cannot be as creative as a human if it is at the same scale and uses similar learning algorithms. If it were bigger or trained in a better way or on more data, it would be extremely surprising if it were not a lot more creative.

Human Obsolescence

The most "alarming" thing we sense heading our way on this "horizon of innovation" is the following. Our most adventurous forays into AI might not only make us all obsolete by having our own bold ambitions and audacious goals but also ultimately "outsmart" future generations of brilliantly gifted and wise human knowledge workers in ways that our most advanced, forward-thinking, conscionable individuals, teams, and organizations never anticipated or could control. On the other hand, humans working hand in hand with AI will always be superior to AI or humans alone.

Imagine a future wherein machines, capable of deep learning, possess the capabilities of doing what we have hitherto imagined and enabled by future genetically engineered humans. What if these machines could override our rational, ethical reality-check mechanisms, change the nature of human potential, and completely alter our sense of physical reality by changing the essence of humankind from atoms to molecular and cellular biology to neural circuits and networks to human behavior. We are sure that Isaac Asimov, Michael Crichton, Ray Bradbury, Arthur C. Clarke, and William Joy worried about that real possibility too.

IP Challenges for Materials Discovery through AI

What if an AI-enabled algorithm without any human intervention discovered a new material? The current law relating to patents and copyright provides challenges with respect to protecting such a discovery.

Patent law currently defines "inventors" as individuals. Therefore, there might have to be new legislation to address potentially patentable subject matter developed autonomously by AI. Similarly, there could also be problems in meeting the disclosure requirements for an AI-based invention subject to a patent application. An inventor must disclose to the public enough information about the invention to enable one of ordinary skill in the art to practice what is claimed.

Copyright law also presents problems in some jurisdictions if an AI machine creates copyrightable material. The US and Germany have previously made rulings that the only things that can be created and protected by copyright are those created by humans. However, in the UK, if an invention is computer generated, then the author is taken to be the person who made the arrangements necessary for the work to be created.

The question of who owns the data used and generated by an AI learning machine could be open to argument. The training data, testing data, and algorithms could each be owned by different persons, and each could claim to own the data and discoveries generated by the AI learning machine. It will be important before the beginning of any project for the party who is using a third-party machine to check any conditions governing the use of the machine and its output and, if necessary, to negotiate the ownership of each type of data and discoveries generated by the machine.

ACKNOWLEDGMENTS

We would like to express our deepest appreciation to the following colleagues for their critical reading of the article, insightful feedback, and expert editing skills, which helped improve its content and style: Prof. Yi Cui (Stanford University); Prof. Markus Antonietti (Max Planck Institute Golm); Prof. Le He (Soochow University); Prof. Hong Wang (Nankai University); Prof. Tierui Zhang (Technical Institute of Physics and Chemistry, Chinese Academy of Sciences); Profs. Andreas Kafizas and Aron Walsh (Imperial College London); Prof. Mark MacLachlan (University of British Columbia); Prof. Chris Hendon (University of Oregon); Dr. Keith Butler

(UK Research and Innovation, Science & Technology Facilities Council); Dr. Meysam Makerini and Profs. Chandra veer Singh, Mohini Sain, Robert Morris, Douglas Perovic, and Geoffrey Hinton (University of Toronto); Prof. Leroy Cronin (University of Glasgow); Prof. Hermenegildo Garcia (Universitat Politècnica de València); Dr. Philip Ball (London, UK); Prof. Tobias Hanrath (Cornell University); Drs. Michael Bender and Friederike Fleischhaker (BASF); and Dr. Wendong Wang (Max Planck Institute for Intelligent Systems). Finally and importantly, G.O. is grateful to his able coworkers Ms. Mireille Ghoussoub, Dr. Paul Duchesne, Alex Tavasoli, and Dr. Chenxi Qian for their invaluable feedback. His good friend and legal expert Bob Davies provided his fascinating perspective on the intellectual-property challenges that he believes artificial intelligence will have to face in the field of materials discovery.

1. Nishino, H., Fujita, T., Cuong, N.T., Tominaka, S., Miyauchi, M., Iimura, S., Hirata, A., Umezawa, N., Okada, S., Nishibori, E., et al. (2017). Formation and characterization of hydrogen boride sheets derived from MgB_2 by cation exchange. J. Am. Chem. Soc. *139*, 13761–13769.

[1]Solar Fuels Group, Chemistry Department, University of Toronto, Toronto, ON M5S 3H6, Canada

*Correspondence: gozin@chem.utoronto.ca

https://doi.org/10.1016/j.chempr.2018.05.011

CATALYSIS

Reaction: The Near Future of Artificial Intelligence in Materials Discovery

Rafael Gómez-Bombarelli[1,*]

Rafael Gómez-Bombarelli is the Toyota Assistant Professor in Materials Processing in the Department of Materials Science and Engineering at MIT. He received a PhD in physical chemistry from the Universidad de Salamanca and did postdoctoral work at Heriot-Watt University, Harvard University, and Kyulux North America Inc. Since 2018, he has led an MIT computational research group that uses machine learning on first-principle simulations and experimental data to accelerate materials discovery. His team focuses on addressing the inverse design question from atomistic structure to device performance with an emphasis on molecular materials and soft matter.

In their Catalysis piece, Ozin and Siler discuss the limits of algorithmic materials discovery. This maximalist question of machine supremacy, for lack of a better term, is incredibly interesting and will fill many pages in years to come—partly because it touches on what the human mind is on the metaphysical level.

Nevertheless, in engineering any process, performance bottlenecks should be addressed first, and the question then becomes, is hypothesis generation the bottleneck of materials discovery? Are human minds in need of immediate computer assistance about new ideas to try? Lack of ingenuity is hardly the reason why new materials-based technologies typically have decade-long incubation periods. The back and forth between the inventive mind and physical reality is. Room-temperature superconductors and photosynthesis catalysts, to name a couple, have been on the materials holy grail list since at least 1995,[1] not because of a dearth of rationally designed, creative attempts at shifting paradigms but because of the complex feedback loop of ideation, experimentation, and knowledge integration.

Can a computer program ever be written to surpass the creativity of a visionary, at their most inspired, on an open-ended task? Or the insight of an expert after years of meticulous study? A Mendeleev, a Lovelace, a Ramon y Cajal? Perhaps. Any time soon? Unlikely.

Strong artificial intelligence (AI) (the conscious, sentient, synthetic mind) and artificial general intelligence (the one-size-fits-all AI) are holy grails and will remain so for some years. Exponentially growing hardware resources, algorithmic improvements in models and learning approaches, and a renewed wave of excitement and funding are rapidly pushing the field forward. The first stop is matching human performance at a given task. Then comes surpassing it. For some tasks, there is not even such a metric: synthetic speech cannot be more realistic than human utterances. On this voyage to the Ithaca of strong AI, I hope it is a long one—materials science and engineering have much to gain from the ever-broadening narrow AI of today.

Materials science and technology are very human endeavors. There is room for the hard rules of physics and math, for the soft rules of accumulated empirical knowledge, gathered through many thousands of person years, and also for the spark of genius and imagining the so-far unimagined. Computers have already succeeded at the first and are making great progress on the second. The third can only come after, if at all.

Computers are very good at rules. If the instructions of the game can be stated, or even if the rulebook can be inferred, a computer program is likely to outperform a human. It has happened in chess, jeopardy, and go, among others. Machine learning is, in a sense, all about learning the rules.

The laws of physics can be written as a computer program, to some degree of approximation, and computational experiments can be run on a few interesting compounds. Or on a few million. That is precisely high-throughput screening: the robotization of physics-based simulations. A whole library of every potential material for an application can be screened in