From Alchemy to AIchemy: On Matter, Minds & Tools

Alán Aspuru-Guzik

This essay explores the impact of information processing on chemistry and materials science through a chronological journey that begins with the origins of life and ends with speculations of what will be possible in 2050. We pause at key moments in this history to reflect on changes to the three main, intertwined threads of this story: matter, minds, and tools. "Matter" refers to the chemical environment that our species interacts with. "Minds" refer to the collective of natural and artificial intelligence that we have access to, as well as our inherent biological information-processing capacity. "Tools" refer to the different processing technologies, such as flasks, oil refineries, computer processors, and so on, that humans employ to understand and manipulate matter. Through this framework, the essay considers the roles of cyborgs and androids today and speculates on the potential future of AI for science.

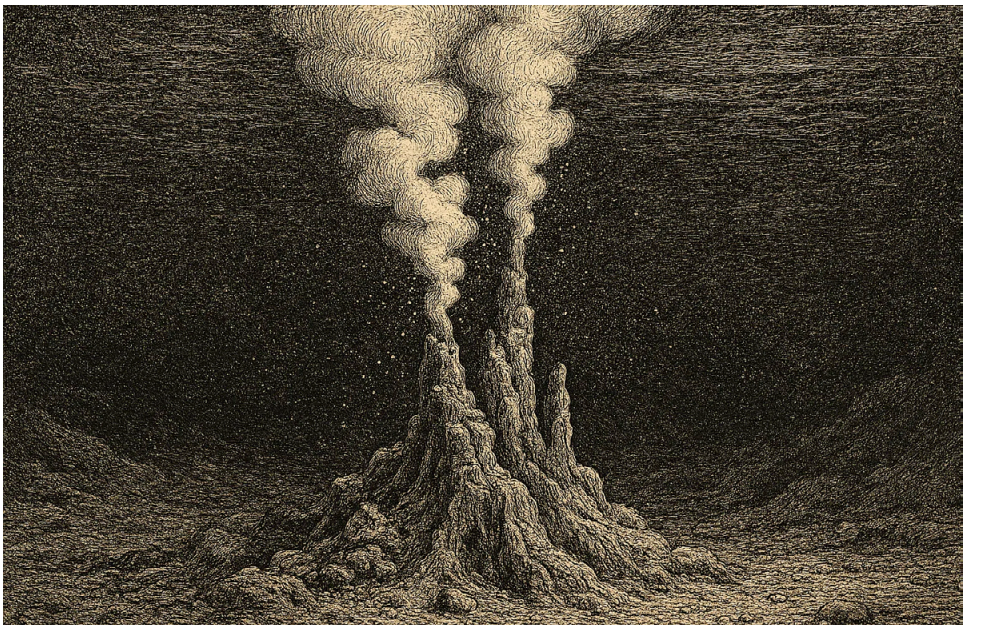
This story begins approximately four billion years ago, when life could have originated at an underwater hydrothermal vent, driven by chemical free-energy gradients generated by reactions between chemical compounds exiting the vent and those in the surrounding ocean (as seen in Figure 1).[1] This heterogeneity created an opportunity for the formation of steady-state chemical reaction networks composed of organic compounds, which, if persistent enough, could have led to the emergence of protolife. This protolife would eventually have learned to encapsulate information within membranes and developed an apparatus for self-replication.

This biological generative process highlights two critical elements of our discussion: *information*, which is related to the complexity of the chemical reaction network itself and the self-replication information that was achieved – this is when life began; and *matter*, which comprises the chemical substrates involved in metabolism and the making of the organism. Information and matter are the two essential components of life.

For most of those four billion years, information was mainly passed down through generations of different life forms by replication, a mechanism known as the fundamental dogma of biology. This process involves ribosomes that transcribe the information stored in deoxyribonucleic acid (DNA) into ribonucleic acid

https://doi.org/10.1162/DAED.a.973

Figure 1
Depiction of an Underwater Hydrothermal Vent



A depiction of a hydrothermal vent, where free-energy gradients could have originated life on Earth. Source : Image generated by the author using prompts in OpenAI's Sora.

(RNA) and then convert it into proteins. Proteins are responsible for the chemical networks that sustain cells, which eventually replicate DNA to produce progeny.

The natural "algorithm" that modifies this genetic information is called *evolution*. In this slow trial-and-error process, the probabilistic nature of replication generates a variety of organisms, some of which become more fit to survive while others become less so. The mutation rate of the genetic information reflects a trade-off between how fast an organism can try out new "solutions" to environmental problems and its own stability.[2] In addition, other forms of information are involved in life, for example, chemical communication between cells. These forms of information exchange were determined primarily by genetic information and therefore evolved at the same pace as evolution. Among bacteria, this phenomenon is known as quorum sensing.[3]

The impact of life on *matter* on Earth was tremendous. The origin of our oxygen-rich atmosphere dates back approximately two billion years to photosynthetic organisms. These early living organisms shaped the entire redox chemistry of the biosphere. To adapt to their environment, organisms grew complex nervous sys-

tems that allowed them to adapt to their environmental circumstances *on the fly*. Let us pause and consider the role of the *mind* in this context. Although the minds of dinosaurs and early mammals were quite sophisticated, and they communicated with each other, the crucial ability to process more complex information and pass it along to future generations was not possible, as far as we know, until the arrival of humans.

Fast-forward to two hundred thousand years ago: we can imagine a group of early hominids sitting around a fire and cooking food while experienced members discuss hunting techniques with younger ones, as shown in Figure 2 (left). Speech was believed to have evolved through a critical mutation in the FOXP2 gene, which may have been a pivotal step in the emergence of complex speech patterns.[4]

In our imagined scene, two key ideas in our discussion are represented: *information* and *matter*. Let us begin with *information*: Through *language*, knowledge and concepts gained from experience can be passed down to new generations in a completely novel way, much faster than through evolution alone. The accumulated information enabled by language facilitated interaction with the environment and other species, which led to rapid changes in prehistoric societies. Equally important in this scene are the chemical transformations that hominids began experimenting with: the use of fire enabled substantial changes in their diets, and the use of tools helped them further transform *matter*, for example, by aiding in the extraction of metals.[5] Tool use, combined with language, therefore allowed our species to change its environment through increasingly complex chemical transformations.

The transformation of matter and the use of novel materials as tools are so essential that we define the ages of humanity by such tool-enabling materials. Here, we focus on the chemical transformations employed in conjunction with such novel materials. For example, Figure 2 (right) showcases a Mesoamerican woman making *masa* from nixtamalized corn. The process of nixtamalization involves soaking and cooking corn in *nextli* (from the Náhuatl word for volcanic or wood ashes, an alkaline source), which increases its nutritional value, improves its flavor, and enables its transformation into *tamalli*, or corn dough, for making tortillas.[6] The *metate* shown in the figure is a stone-made tool that allows her to grind the nixtamalized corn and eventually flatten the masa. The masa is then cooked over fire to make tortillas. All these steps involve acid-base chemistry and fire to create one of the main staples of Mesoamerican civilizations.

Over the past two and a half million years, humanity has progressed from the Stone Age to the Bronze Age to the Iron Age and, within what is known as the Anthropocene, to the plastics age, the silicon age, and perhaps the era of materials by design.[7] The written and archaeological records from these periods have revealed an intricate history of chemical transformations. For example, throughout history, humanity has utilized dyes to adorn clothing, decorate living

Figure 2
Depictions of Early Hominids and Humans Processing and Sharing Information

(Left) Hominids cook and talk around a fire. Two main themes of this essay are depicted in this image : first, the use of language to transfer information among generations ; second, our use of fire to transform food and begin the era of human-induced chemical transformations. *(Right)* A Mesoamerican woman using a *metate* stone tool to prepare *masa* (corn dough). The *masa* had been previously chemically transformed by the nixtamalization process. These AI-generated images are meant to illustrate the intimate connection between cooking, chemistry, and our own development as a species. Source : Images generated by the author using prompts in OpenAI's Sora.

spaces and structures, and create a wide range of art forms. The ancient production of dyes and pigments originated from a complex set of biological and mineral sources.[8] For instance, indigo is a plant-extracted dye used by several civilizations, from the Incas to the Egyptians.[9] We now produce indigo synthetically, mostly to make our blue jeans blue.

For thousands of years, such chemical transformations were mainly empirical, and when they were not, they were described in terms of the philosophy of nature at the time. For example, to explain the origins of life, seventeenth-century Jesuit scholar and polymath Athanasius Kircher attributed them to a primordial "salino-sulphur-mercurial vapour," which was the universal seed for not only inanimate but also animate objects.[10] Half a century younger than Kircher, Isaac Newton was also an avid alchemist and devoted considerable time to his quest for the elixir of life and the philosopher's stone, which was believed to transform common metals into gold.[11]

The transition into modern chemistry began around this time with scientists such as Robert Boyle and Robert Hooke, who were inspired by the emerging inductive philosophy of Francis Bacon, laying the groundwork for our current, systematic, and scientific approach to chemical problems.[12] One century later, in 1789, the use of induction and deduction led to the critical experiments of Henry Cavendish and Antoine Lavoisier, which dethroned the phlogiston theory and led to the discovery

of nitrogen and oxygen, respectively. The publication of Lavoisier's *Traité élémentaire de chimie* was as revolutionary as the French Revolution, which occurred the same year as its publication.[13]

We will skip the next couple of centuries. Although the field of chemistry has advanced substantially, including the development of chemical bond theories and the emergence of quantum chemistry, the way humans process information, by passing down their discoveries to new generations in verbal and written forms, has remained essentially unchanged until recently.[14] The Industrial Revolution, in part, led to the creation of the bureaucratic state.[15] The creation of the Computing-Tabulating-Recording-Company, now known as International Business Machines (IBM), was born out of the necessity of having an accurate population census. This, coupled with the need for extensive computing and tabulating operations during the Second World War, led to the birth of what we would now know as modern computing and eventually computer science.[16] In what follows, we will intertwine the histories of chemistry and materials science with that of computer science. We will only briefly discuss the history of computing as it pertains to our story, focusing on its back-and-forth influence with chemistry and materials science. We will make stops at critical decades – the 1950s, the 1980s, the 2010s, and the current 2020s – and end the essay speculating about the intersection of these two fields in 2050.

During the twentieth century, the various subdisciplines, including organic, inorganic, and analytical chemistry, have established robust foundations and developed a rational framework for the understanding of chemicals and materials. For example, the pioneering work of chemist Christopher Ingold in the 1920s led to the concept of organic reaction mechanisms, which were designed to rationalize the complex behavior of matter. At Harvard University, organic chemist Robert Burns Woodward was a master at designing synthetic routes to increasingly complex natural products, accomplishing the synthesis of complex molecules such as cortisone, an immunosuppressant, in 1951, and reserpine, an antihypertensive, in 1956.[17]

The 1950s marked a pivotal period in the advancement of computational chemistry, as computers became increasingly accessible to scientists. In the same year that reserpine was synthesized, theoretical chemist Samuel Francis Boys and coworkers performed the first general quantum-chemical calculation on a modern computer.[18] In 1957, physicists Berni Alder and Thomas Wainwright developed the first molecular dynamics code to study the physics of hard-sphere systems.[19] The same year, Louis C. Ray and Russell A. Kirsch initiated the first cheminformatics database of chemical structures.[20] A couple of years later, theoretical chemists Christopher Longuet-Higgins and Lionel Salem investigated the alternating nature of chemical bonds in ethylene.[21] These are just a few examples of the significant breakthroughs and publications that marked the beginning of the era of modern computer infor-

mation and chemistry (cheminformatics), as well as the interface of simulation and chemistry (computational and theoretical chemistry).

The Dartmouth Summer Research Project on Artificial Intelligence, a founding event in the field of AI, also began serious discussions in the 1950s to refine and advance ideas about thinking machines.[22] At the time, deliberations were centered around cybernetics, a nascent field spearheaded by physiologist Arturo Rosenblueth and mathematician Norbert Wiener, whose concepts, such as feedback mechanisms and self-organization in complex systems, are explored in this issue of *Dædalus*.[23]

Artificial intelligence and its subfields are highly relevant to the thread of this essay, as they represent human attempts to build machinery that can replicate human thought processes. This means that if the field is successful in generating powerful enough systems, information can now be processed, transformed, and communicated from AIs to other AIs as well as to humans. This would mark the beginning of the era of *androids* or artificial thinking beings. Even if partially successful, AI promises several societal changes. It is not surprising that, despite such great promise, the significant mismatch between the power of the computers available to researchers and researchers' algorithmic understanding of how to build intelligent systems has led the field to undergo several periods of excitement (springs) followed by disillusionment (winters).

This story now stops in the 1980s. With the advent of the microprocessor, computing became even more ubiquitous in science, thereby profoundly impacting chemistry. In the basement of Harvard's chemistry department, several scientists were working with organic chemist Elias James Corey on the LHASA (Logic and Heuristics Applied to Synthetic Analysis) computer program, a computational formalization of the concept of *retrosynthesis*: the idea that the (total) synthesis of a molecule can be broken down retroactively into the original fragments.[24] The computational "game" of retrosynthesis is not so dissimilar to the game of chess. This systematic integration of information and chemistry ultimately led to Corey receiving the 1990 Nobel Prize in Chemistry for his development of the theory and methodology of organic synthesis.[25]

In the same decade, researchers made significant progress in the field of artificial intelligence. A former graduate student of Longuet-Higgins, computer scientist Geoffrey Hinton was studying Boltzmann machines and developing the powerful backpropagation approach that led to the modern theory of neural networks.[26] This work ultimately led to him receiving both the 2018 Turing Award (together with computer scientists Yoshua Bengio and Yann LeCun) for work on deep learning, and the 2024 Nobel Prize in Physics (with physicist John Hopfield) "for foundational discoveries and inventions that enable machine learning with artificial neural networks."[27] This Nobel Prize–winning work is an example of the formal connections between the theory of matter (statistical mechanics) and those of machine learning and artificial intelligence, which may be why in 1967 Longuet-Higgins switched

fields from theoretical chemistry to learning and cognition, where he made seminal contributions and also supervised Hinton. Unfortunately, although the decade of the 1980s saw huge progress, the results obtained during that period were not yet ready for prime time, as the data and computing capabilities were still not up to the requirements of the developed theory. This led to an AI winter that lasted for about two decades.

In the 2010s, another computational revolution began with the advent of the general-purpose graphics processing unit (GPGPU, or GPU for short) introduced by NVIDIA, which enabled scientists to perform more complex calculations on desktop computers. In 2012, at the University of Toronto, Hinton's team developed AlexNet, a neural network that became the state of the art for image recognition on the ImageNet dataset.[28] Around that time, there was a significant resurgence of interest in neural networks, including the hiring of AI researchers at top universities. For example, Harvard hired computer scientist Ryan P. Adams, a former postdoctoral researcher with Hinton, who cofounded the Whetlab machine learning startup while an assistant professor at the university from 2011 to 2016. I was also a faculty member at Harvard from 2006 to 2018, where I was lucky enough to participate in early research on AI for chemistry, as well as to engage with Longuet-Higgins's two career interests: *mind* and *matter*. Incidentally, I was trained as a computational chemist with an academic lineage that traced back to theoretical chemist John Pople, winner of a Nobel Prize in Chemistry and a collaborator of Longuet-Higgins.[29]

After Adams arrived at Harvard, his interactions with students like David Duvenaud (a postdoc in his research group) and Rafael Gómez-Bombarelli (a postdoctoral fellow in my lab) – along with Adams's interactions with myself and many others outside of our groups – led to several key developments in the field of machine learning for chemistry. In particular, we developed a graph convolutional neural network for predicting molecular properties, which we successfully employed to discover organic light-emitting materials.[30] Our publication of this innovative network helped attract several machine learning researchers to the field of chemistry. Ten years later, it is common to see chemistry and materials science publications at the major computer science conferences such as the Conference on Neural Information Processing Systems (NeurIPS), the International Conference on Learning Representations (ICLR), and the International Conference on Machine Learning (ICML). Coincidentally, the room at Harvard where these publications originated as research, and which was also part of my research laboratory, is the same basement room that was once occupied by the students of Elias Corey, who developed the LHASA program forty years prior.

In 2016, we developed a generative model for the inverse design of molecules.[31] Generative models are a potent and "magical" form of AI. Given a complex distribution, such as all the images or text on the internet or several voice recordings,

these models generate artificial samples from the distribution. In other words, the model can produce text, images, or voice recordings that did not exist in the training dataset yet are statistically similar to it. So far, it seems useful, but *not as useful* as when we can condition such a complex distribution on a desired set of attributes. For example, one could ask for audio samples that sound like the legendary singer Yma Sumac performing a song by José Alfredo Jiménez. Solving these types of problems is sometimes called *inverse design*, as one wants to create, from a distribution, novel samples that satisfy certain design conditions. This is in contrast to most computational methods, which employ *direct simulation* that links samples to their simulated properties. Inverse design opens the door to searching for (and finding) needles in a haystack.

Generative models for chemistry are extremely useful, given the enormous chemical space: It is, in principle, infinite, yet estimations of synthesizable, practical molecules are up to a range from 1,060 to 10,180. So the enumeration of these species is practically impossible using traditional methods. Here is where generative models shine.

Another key encounter during this part of my career at Harvard was with Alex Zhavoronkov, the CEO and cofounder of InSilico Medicine. At the time, Alex and his early-stage startup company were also experimenting with some of the earliest generative models in chemistry. Together, we established a collaboration that demonstrated the ability to generate a lead candidate for inhibiting the DDR1 kinase within forty-five days.[32] This study and subsequent work laid the foundation for AI-based drug discovery. As of today, practically all the large pharma companies have significant AI drug-discovery programs.

Ten years ago, it was very common to say machine learning *tools*, as many of us didn't think of these models as *minds*. If we take the position that AI, even as of ten years ago, is a form of mind, it would still require physical tools to transform *matter*: even if it is possible for an AI system to suggest a chemical reaction for a human to carry out, it would be much more effective if the AI were empowered to carry out the reaction itself by means of automated equipment, analyze the results, and use its data for further refinement of a hypothesis or optimization of a desirable property. A significant number of scientists had already been thinking along these lines. In 2017, under the auspices of the global Mission Innovation initiative, which was spawned by the 2015 Paris Agreement on climate change, physicist and chemist Kristin Persson and I organized a conference in Mexico that led to a report that coined the term "materials acceleration platform," now also known as a self-driving laboratory.[33] This conference included the participation of several pioneers of the field – such as Marty Burke, Jason Hein, Benji Maruyama, Tejs Vegge, and many others – and was the self-driving laboratory analogous moment to the original Dartmouth conference on artificial intelligence.

The concept of a self-driving lab represents a significant milestone. So far, the *human mind* has been the primary factor influencing the transformations of *matter*. However, as self-driving laboratories become increasingly intelligent and autonomous, these robotic platforms and devices have become an extension of humans, turning them into *cyborgs* – that is, humans augmented by machines – that conduct science.[34] Recent discoveries from self-driving labs include new conductive thin-film formulations, organic lasers, and perovskite nanocrystals.[35]

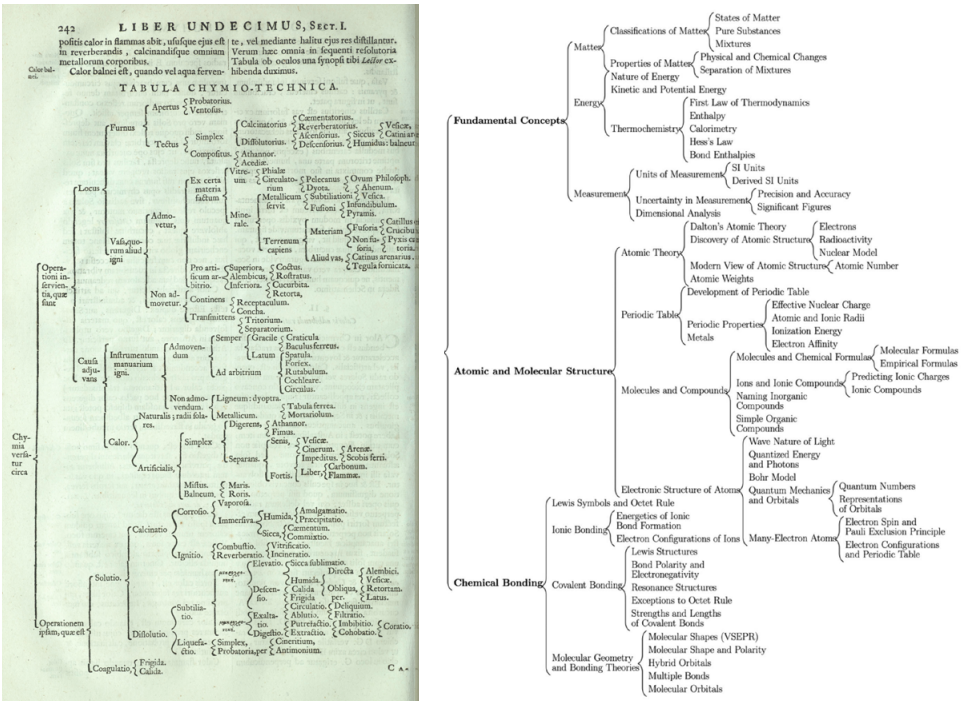
Back in Boston in 2018, inspired by the themes of the time, chemists Markus Reiher, Roland Lindh, and I published a manifesto outlining the new directions of research in computational chemistry for the twenty-first century, which we called "(R)evolution."[36] In that work, we argue that theoretical chemistry in this century is closely tied to the technological revolutions of our time. We imagined a dialogue between a self-driving laboratory called Organa and a chemist named Jane. In this dialogue, the lab would interact with the chemist and help plan and execute an experiment. Confronting this challenge would require decades of community research. Nevertheless, just seven years later, Organa exists, and it performs experiments at our laboratory in Toronto (*vide infra*).

It is time to move forward to the current decade, the 2020s. This is the second issue of *Dædalus* this decade concerning AI, which is no coincidence given the current global focus on artificial intelligence and its implications for society.[37] Generative models of text and images have become ubiquitous and are predicted to transform almost every industry. The advent of large language models (LLMs) has transformed what is possible to achieve in AI at scale.[38] As an example of the application of LLMs in chemistry, let us compare summaries of the fields of alchemy and modern chemistry, as seen from the perspective of an undergraduate textbook. In Figure 3, we present (left) the classification of alchemy topics described by Athanasius Kircher in his 1665 scientific textbook *Mundus Subterraneus* and (right) an AI-generated summary tree built from a general chemistry textbook using TreeReader.[39]

Now things come full circle. Athanasius Kircher spent a lifetime collecting what was known in the world for his "textbook" *Mundus Subterraneus*. He then further synthesized his knowledge into the tree-like table shown in Figure 3. Our tool TreeReader spent less than a minute summarizing a modern chemistry textbook. The striking side-by-side comparison of the differences and similarities, not only of how the trees were generated (human versus AI) but also how the field of chemistry has changed (alchemy versus the scientific method), refer deeply to the threads in this essay.

In a short time since the advent of LLMs, voice- and vision-controlled self-driving labs have become feasible. Based on these developments, and in collaboration with computer scientists Florian Shkurti and Animesh Garg, we have realized our dream

Figure 3
Alchemy and Modern Chemistry Compared Using Summary Trees



(Left) The classification of alchemical concepts as described by Athanasius Kircher in *Mundus Subterraneus, in XII libros digestus* (1665). *(Right)* An AI-derived summary of the classification of chemical concepts in a modern chemistry textbook. Source : Athanasius Kircher, *Mundus Subterraneus, in XII libros digestus* (Johannes Janssonius and Elizeus Weyerstraten, 1665), https ://dpul .princeton.edu/alchemy/catalog /2227mw62t ; and Zijian Zhang, Pan Chen, Fangshi Du, et al., "TreeReader : A Hierarchical Academic Paper Reader Powered by Language Models," arXiv (2025), https ://doi.org/10.48550/arXiv.2507.18945.

of developing and demonstrating Organa. We built Organa as an agentic assistant that can communicate with an end user to help plan electrochemistry experiments, surpassing our initial prediction of how long it would take to carry out this task![40] Organa belongs to the large class of systems known as self-driving labs and is also part of a vast field of agents that conduct scientific research.[41]

A recent review highlights the vast array of agentic systems for science.[42] Most of those agents are *minds* without physical *tools*. They utilize digital tools to conduct literature reviews, formulate hypotheses, and perform computational experiments. After developing Organa, we introduced El Agente, a computational chemistry agent

developed by over thirty people, as described by our graduate student Yunheng Zou and colleagues in "El Agente: An Autonomous Agent for Quantum Chemistry."[43] As of today, El Agente can solve computational chemistry tasks at the level of an advanced undergraduate student. Connecting El Agente to self-driving labs is a current and meticulous work in progress, as one must carefully consider all possible safety guardrails.[44]

Today, agents for science are operating in ways that are already transforming research, and they are on the verge of evolving into forms we can barely imagine. This is because large companies such as Anthropic, OpenAI, NVIDIA, and Google are continually adding base agentic features to their models, making them more capable of performing complex research tasks.

First, we can think of a straightforward *shallow* agent as a contextual description (which can be thought of as a "pre-prompt" or a general set of guidelines provided to a language model), followed by a set of instructions on how to approach a particular task or use a given tool. These agents could be valuable for mundane extract, transform, and load (ETL) operations, which remain among the most sought-after agent workflows, as data from academic groups and large companies can be convoluted in legacy documents that are not easily parsed. The notation of such papers can also be problematic. An example of a shallow agent we built in our Toronto lab is MERMaid, which extracts reaction conditions from legacy scientific PDF documents.[45] MERMaid builds a knowledge graph to store the extracted information, enabling it to construct a scientific world model across different publications.

For more complex tasks, such as generating hypotheses or focusing on the use of different tools, El Agente is considered a *deep* agent.[46] By that, I mean it is built on a hierarchical cognitive architecture inspired by frameworks like Soar and CoALA, designed to provide LLM agents with structured reasoning and long-term autonomy.[47] At its core are several interlinked types of memory. The working memory combines four elements: 1) a global memory that stores shared user directives and molecular details, ensuring all agents maintain a coherent context; 2) agent-specific conversation histories that record local decisions and dialogues; 3) a grounding mechanism that links agents' reasoning to the actual computational environment (such as file structures); and 4) a long-term memory divided into procedural, semantic, and episodic layers. Procedural memory encodes how agents collaborate to perform tasks, semantic memory stores patterns and domain knowledge, and episodic memory retains past experiences to facilitate continual learning. Together, these memory systems enable the architecture to function in a manner that mirrors cognitive processes – planning, reasoning, and self-critique – across multiple timescales.

This architecture functions as a hierarchy of agents working in concert. At the top sits the computational chemistry agent, responsible for high-level planning and interpreting the user's natural-language goals. Beneath it, mid-level agents

manage specialized domains – such as geometry generation and optimization, quantum calculations, and file management – analogous to departments in a complex organization. Each of these mid-level agents oversees lower-level expert agents that execute concrete actions, such as generating quantum chemistry input files, submitting jobs to high-performance computing clusters, or diagnosing errors. Communication flows downward through task delegation and upward through feedback and summarized reporting, allowing the system to maintain clarity even in multistep workflows. This cybernetic network design strikes a balance between autonomy and coordination : each agent acts locally, using its own contextual memory and reasoning core, yet collectively they form a coherent, adaptive whole. The result is a system capable of long-term reasoning, adaptive error correction, and transparent action logging, making the scientific process both traceable and intelligible to human collaborators.

Let's pause for a moment to further discuss the tools that El Agente employs. These quantum chemistry and molecular simulation tools are approximate *digital twins* of the quantum world. They solve the Schrödinger equation numerically, using an approximate method. The process of employing these programs is quite intricate. Therefore, a deep agentic system is needed to describe the input and parse the output of these programs, which comprise up to millions of lines of numerical code. Incidentally, these codes are the descendants of the computer packages that Samuel Boys developed in the 1950s.[48] For a medium-sized molecule, these calculations can consume hours of computer time on a modern computer cluster. Therefore, agents for science can end up orchestrating complex and costly computational workflows.

To make El Agente more closely resemble a scientist, we need to add hypothesis-generation capabilities to the tool. An ongoing project that bridges this gap is called PaperMon, a dual homage to the demons of science – imagined beings summoned in thought experiments to challenge scientific theories, such as Maxwell's demon – and Pokémon.[49] PaperMon enables the instantiation of a small demon per scientific paper. These agents have contextual information about a particular scientific paper or set of documents and become advocates for them. We then built a system that enables PaperMon to discuss with each other and generate sensible hypotheses. There is a significant gap between what we do nowadays with these tools and the work of a full-blown AI scientist. Still, the path between them is clear : the artificial scientist needs to be able to *generalize* to different problems and to clearly and understandably describe the scientific concepts it understands to others.[50]

During the months I have spent writing this essay, large-scale advances in venture capital investment in AI and materials science have sparked renewed global interest in their intersection. Companies like Lila Sciences, Periodic Labs, Orbital Materials, and CuspAI have received more investment than the entire annual budget of the U.S. National Science Foundation in the fields of

chemistry and materials science.[51] The United States has also announced the Genesis Mission, a project being compared in scope to the Manhattan Project, to advance AI for science, with materials as a primary focus.[52] The year 2026 is when all these threads come together, attracting both private and public interest.

Making predictions is a risky business, but I would like to consider what the intersection of artificial intelligence and chemistry will look like in 2050. We may have already developed AI systems that match or surpass a human's ability to generate hypotheses and design experiments. Furthermore, the embodiment of these AIs would enable them to conduct the experiments themselves, thereby creating fully autonomous, self-driving laboratories. These systems would not be *cyborgs* or human-machine hybrids but rather full-blown *androids* or AI scientists.

As we conclude our chronological journey, it is worthwhile to revisit the geophysical impact of the evolution of oxygen-evolving organisms. They entirely transformed our biosphere. It is also worthwhile to speculate and reflect on potential implications for our environment from these large-scale AI scientists. Data center energy consumption is expected to account for a significant fraction of global energy demand by 2050. Suppose AI scientists succeed in undertaking noble scientific and engineering endeavors. In that case, they will have contributed to humanity's transition to a green economy by discovering new materials and engineering them into sustainable energy systems that align with the survival of our ecosystems.

On the other hand, if the current "business as usual" trajectory holds, we may need to devise a new term for the era that could follow the Anthropocene. For example, the *AIocene* would be an era in which AI itself is responsible for large-scale changes to the biosphere, potentially and irreversibly altering ecosystems. Here, what started as a computational *tool* would have become a *mind* of its own, capable of using its own tools (self-driving labs) and affecting *matter* at a planetary scale through its own energy consumption. And this is how our story would come full circle, once again.

AUTHOR'S NOTE

Most of the images in this essay were generated as a *cyborg* between simple prompts and OpenAI's text-to-video model Sora. They are inspired by book illustrations from the 1800s, which in turn take me back to the readings of several Jules Verne books I devoured cover to cover when I was around ten years old. I dedicate this work to Varinia Bernales for the valuable discussions that inspired several aspects of this manuscript, as well as her support and edits that helped bring it to completion. I would also like to congratulate Zijian Zhang for generating Figure 3 (right) using his tool TreeReader. Finally, I would like to express my gratitude to Dr. Anders G. Frøseth for his continued support of my scholarship and his friendship.

ABOUT THE AUTHOR

Alán Aspuru-Guzik is Director of the Acceleration Consortium and Professor of Chemistry and Computer Science at the University of Toronto. He is also a Senior Director of Quantum Chemistry at NVIDIA. He has published in such journals as *Science*, *Nature*, and *Proceedings of the National Academy of Sciences*.

ENDNOTES

[1] William Martin, John Baross, Deborah Kelley, and Michael J. Russell, "Hydrothermal Vents and the Origin of Life," *Nature Reviews Microbiology* 6 (11) (2008): 805–814.

[2] Dalit Engelhardt and Eugene I. Shakhnovich, "Mutation Rate Variability as a Driving Force in Adaptive Evolution," *Physical Review E* 99 (2019): 022424.

[3] Michiko E. Taga and Bonnie L. Bassler, "Chemical Communication Among Bacteria," *Proceedings of the National Academy of Sciences* 100 (S2) (2003): 14549–14554.

[4] Michael P. Richards, "A Brief Review of the Archaeological Evidence for Palaeolithic and Neolithic Subsistence," *European Journal of Clinical Nutrition* 56 (2002): 1262–1278; and Rafael Mier, *Maíz: Origen, cultura y cocina* (Ediciones Larousse, 2025).

[5] Richards, "A Brief Review of the Archaeological Evidence for Palaeolithic and Neolithic Subsistence."

[6] Mier, *Maíz*.

[7] Colin N. Waters, Jan Zalasiewicz, Colin Summerhayes, et al., "The Anthropocene Is Functionally and Stratigraphically Distinct from the Holocene," *Science* 351 (6269) (2016), https://doi.org/10.1126/science.aad2622; Jennifer A. Brandon, William Jones, and Mark D. Ohman, "Multidecadal Increase in Plastic Particles in Coastal Ocean Sediments," *Science Advances* 5 (9) (2019), https://doi.org/10.1126/sciadv.aax0587; and Troy Scott, Amanda Walsh, Benjamin Anderson, et al., "High-Tech Infrastructure and Economic Growth: The Materials Genome Initiative," *Science and Public Policy* 48 (5) (2021): 649–661.

[8] Omar Abdel-Kareem, "History of Dyes Used in Different Historical Periods of Egypt," *Research Journal of Textile and Apparel* 16 (4) (2012): 79–92.

[9] Jeffrey C. Splitstoser, Tom D. Dillehay, Jan Wouters, and Ana Claro, "Early Pre-Hispanic Use of Indigo Blue in Peru," *Science Advances* 2 (9) (2016), https://doi.org/10.1126/sciadv.1501623.

[10] Daniel Stolzenberg, ed., *The Great Art of Knowing: The Baroque Encyclopedia of Athanasius Kircher* (Stanford University Libraries, 2001); and Hiro Hirari, "Kircher's Chymical Interpretation of the Creation and Spontaneous Generation," in *Chymists and Chymistry: Studies in the History of Alchemy and Early Modern Chemistry*, ed. Lawerence M. Principe (Science History Publications, 2007), 77–87, https://backoffice.biblio.ugent.be/download/376797/6797072.

[11] Richard S. Westfall, *Never at Rest: A Biography of Isaac Newton* (Cambridge University Press, 1983).

[12] John Read, *From Alchemy to Chemistry* (Dover Publications, 2003).

[13] Antoine Lavoisier, *Traité élémentaire de chimie* (1789).

[14] Kostas Gavroglu and Ana Simões, *Neither Physics nor Chemistry: A History of Quantum Chemistry* (MIT Press, 2011).

[15] David Graeber, *The Utopia of Rules: On Technology, Stupidity, and the Secret Joys of Bureaucracy* (Melville House, 2015).

[16] Peter Galison, *Image and Logic: A Material Culture of Microphysics* (University of Chicago Press, 1997).

[17] Ibid.; and R. B. Woodward, F. E. Bader, H. Bickel, et al., "The Total Synthesis of Reserpine," *Journal of the American Chemical Society* 78 (9) (1956): 2023–2025.

[18] Woodward, Bader, Bickel, et al., "The Total Synthesis of Reserpine."

[19] Berni J. Alder and Thomas E. Wainwright, "Phase Transition for a Hard Sphere System," *The Journal of Chemical Physics* 27 (5) (1957): 1208–1209.

[20] Louis C. Ray and Russell A. Kirsch, "Finding Chemical Records by Digital Computers," *Science* 126 (3278) (1957): 814–819.

[21] Hugh Christopher Longuet-Higgins and Lionel Salem, "The Alternation of Bond Lengths in Long Conjugated Chain Molecules," *Proceedings of the Royal Society A* 251 (1265) (1959): 172–185.

[22] Ronald Kline, "Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence," *IEEE Annals of the History of Computing* 33 (4) (2011): 5–16.

[23] Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow, "Behavior, Purpose and Teleology," *Philosophy of Science* 10 (1) (1943): 18–24.

[24] Elias J. Corey, Alan K. Long, and Stewart D. Rubenstein, "Computer-Assisted Analysis in Organic Synthesis," *Science* 228 (4698) (1985): 408–418.

[25] The Nobel Foundation, "Nobel Prize in Chemistry 1990," https://www.nobelprize.org/prizes/chemistry/1990/summary (accessed February 17, 2026).

[26] David H. Ackley, Geoffrey E. Hinton, and Terrence J. Sejnowski, "A Learning Algorithm for Boltzmann Machines," *Cognitive Science* 9 (1) (1985): 147–169; and Werner Kutzelnigg, "H. C. Longuet-Higgins and J. A. Pople: Theoretical Chemistry," *Angewandte Chemie International Edition* 43 (21) (2004): 2740–2743.

[27] Association for Computing Machinery, "Fathers of the Deep Learning Revolution Receive ACM A. M. Turing Award," https://awards.acm.org/about/2018-turing (accessed February 17, 2026); and The Nobel Foundation, "The Nobel Prize in Physics 2024," https://www.nobelprize.org/prizes/physics/2024/press-release (accessed February 17, 2026).

[28] David Duvenaud, Dougal Maclaurin, Jorge Aguilera-Iparraguirre, et al., "Convolutional Networks on Graphs for Learning Molecular Fingerprints," in *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, ed. Corinna Cortes, Neil Lawrence, Daniel Lee, et al. (Curran Associates, Inc., 2015), https://www.proceedings.com/content/030/030770webtoc.pdf.

[29] Kutzelnigg, "H. C. Longuet-Higgins and J. A. Pople"; and The Nobel Foundation, "John Pople–Facts," https://www.nobelprize.org/prizes/chemistry/1998/pople/facts (accessed February 17, 2026).

[30] Duvenaud, Maclaurin, Aguilera-Iparraguirre, et al., "Convolutional Networks on Graphs for Learning Molecular Fingerprints"; and Alex Zhavoronkov, Yan A. Ivanenkov, Alex Aliper, et al., "Deep Learning Enables Rapid Identification of Potent DDR1 Kinase Inhibitors," *Nature Biotechnology* 37 (9) (2019): 1038–1040.

[31] Rafael Gómez-Bombarelli, Jennifer N. Wei, David Duvenaud, et al., "Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules," *ACS Central Science* 4 (2) (2018) : 268–276 ; and Benjamin Sanchez-Lengeling and Alán Aspuru-Guzik, "Inverse Molecular Design Using Machine Learning : Generative Models for Matter Engineering," *Science* 361 (6400) (2018) : 360–365.

[32] Zhavoronkov, Ivanenkov, Aliper, et al., "Deep Learning Enables Rapid Identification of Potent DDR1 Kinase Inhibitors."

[33] Mission Innovation, "Press Release–Materials Workshop Report," January 26, 2018, https ://mission-innovation.net/press-release-materials-workshop-report ; and Gary Tom, Stefan P. Schmid, Sterling G. Baird, et al., "Self-Driving Laboratories for Chemistry and Materials Science," *Chemical Reviews* 124 (16) (2024) : 9633–9732.

[34] Benjamin P. MacLeod, Fraser G. L. Parlane, Thomas D. Morrissey, et al., "Self-Driving Laboratory for Accelerated Discovery of Thin-Film Materials," *Science Advances* 6 (20) (2020), https ://doi.org/10.1126/sciadv.aaz8867.

[35] Ibid. ; Felix Strieth-Kalthoff, Han Hao, Vandana Rathore, et al., "Delocalized, Asynchronous, Closed-Loop Discovery of Organic Laser Emitters," *Science* 384 (6697) (2024), https ://doi.org/10.1126/science.adk9227 ; and Sina Sadeghi, Karl Mattsson, Joshua Glasheen, et al., "A Self-Driving Fluidic Lab for Data-Driven Synthesis of Lead-Free Perovskite Nanocrystals," *Digital Discovery* 4 (7) (2025) : 1722–1733.

[36] Alán Aspuru-Guzik, Roland Lindh, and Markus Reiher, "The Matter Simulation (R)evolution," *ACS Central Science* 4 (2) (2018) : 144–152.

[37] James Manyika, "Getting AI Right : Introductory Notes on AI & Society," *Dædalus* 151 (2) (Spring 2022) : 5–27, https ://www.amacad.org/publication/daedalus/getting-ai-right-introductory-notes-ai-society.

[38] Ashish Vaswani, Noam Shazeer, Niki Parmar, et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, ed. Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, et al. (Curran Associates, Inc., 2017).

[39] Jiaqi Wei, Yuejin Yang, Xiang Zhang, et al., "From AI for Science to Agentic Science : A Survey on Autonomous Scientific Discovery," arXiv (2025), http ://arxiv.org/abs/2508.14111 ; Yunheng Zou, Austin H. Cheng, Abdulrahman Aldossary, et al., "El Agente : An Autonomous Agent for Quantum Chemistry," *Matter* 8 (7) (2025), https ://doi.org/10.1016/j.matt.2025.102263 ; Athanasius Kircher, *Mundus Subterraneus, in XII libros digestus* (Johannes Janssonius and Elizeus Weyerstraten, 1665), https ://dpul.princeton.edu/alchemy/catalog/2227mw62t ; and Theodore E. Brown, H. Eugene LeMay, Bruce E. Bursten, et al., *Chemistry : The Central Science*, 12th ed. (Pearson Education, 2013).

[40] Theodore R. Sumers, Shunyu Yao, Karthik Narasimhan, and Thomas Griffiths, "Cognitive Architectures for Language Agents," *Transactions on Machine Learning Research* (2024), https ://openreview.net/pdf ?id=1i6ZCvflQJ.

[41] Tom, Schmid, Baird, et al., "Self-Driving Laboratories for Chemistry and Materials Science."

[42] Wei, Yang, Zhang, et al., "From AI for Science to Agentic Science."

[43] Zou, Cheng, Aldossary, et al., "El Agente."

[44] Sumers, Yao, Narasimhan, and Griffiths, "Cognitive Architectures for Language Agents."

[45] John E. Laird, *The Soar Cognitive Architecture* (MIT Press, 2021).

[46] Zou, Cheng, Aldossary, et al., "El Agente."

[47] Laird, *The Soar Cognitive Architecture*; and Sumers, Yao, Narasimhan, and Griffiths, "Cognitive Architectures for Language Agents."

[48] Samuel F. Boys, Gregory B. Cook, Colin M. Reeves, and Isaiah Shavitt, "Automatic Fundamental Calculations of Molecular Structure," *Nature* 178 (4544) (1956): 1207–1209.

[49] Mario Krenn, Robert Pollice, Si Yue Guo, et al., "On Scientific Understanding with Artificial Intelligence," *Nature Reviews Physics* 4 (12) (2022): 761–769.

[50] Ibid.

[51] Charles Yang, "Venture Capital is Subsidizing U.S. Material Science Research," ML4Sci, October 15, 2025, https://ml4sci.substack.com/p/venture-capital-is-subsidizing-us.

[52] The White House, "Launching the Genesis Mission," November 24, 2025, https://www.whitehouse.gov/presidential-actions/2025/11/launching-the-genesis-mission.