https://cen.acs.org/business/informatics/lab-future-ai-automated-synthesis/99/i11



INFORMATICS

# The lab of the future is now

## Recent demonstrations of AI-directed automation may herald a new world for drug and materials discovery

*by* ***Rick Mullin***

MARCH 28, 2021 | APPEARED IN **VOLUME 99, ISSUE 11**




**IN BRIEF**

**After making gradual inroads in drug and materials discovery in recent years, artificial intelligence is suddenly appearing in a new light.** Autonomous labs that link AI-based data analysis to robotic synthesis and validation are being demonstrated in both academia and industry. These labs prefigure a new work environment in which machines perform many of the traditional

High-tech labs unveiled by academic researchers, a computing giant, and a major drug company all blend artificial intelligence computing and robotics in ways that may herald a new world of research. They suggest a kind of Renaissance lab for multidisciplinary scientists steeped in chemistry, biology, and data science.

Alán Aspuru-Guzik, a professor of chemistry and computer science at the University of Toronto, and colleagues **reported in Science Advances the discovery** of thin-film materials in a “self-driving

**MOST POPULAR IN BUSINESS**

**Can mRNA disrupt the drug industry?**

**Pfizer, Moderna ready vaccine manufacturing networks**

**Is ammonia the fuel of the future?**

**Can superglues get better? Arkema thinks so.**

laboratory" in which AI controls automated synthesis and validation in a cycle of machine-learning data analysis. Andrew I. Cooper, director of the Materials Innovation Factory at the University of Liverpool, and colleagues **published results** from an AI-directed robotics lab that optimized a photocatalytic process for generating hydrogen from water after running about 700 experiments in 8 days (*Nature* 2020, DOI: **10.1038/s41586-020-2442-2**).

functions of researchers. Proponents argue that this self-driven R&D environment will enhance the efforts of multidisciplinary scientists, who will maintain ultimate control. Chemical and drug firms are taking a serious look at what may be the lab of the future.

**Lipids, the unsung COVID-19 vaccine component, get investment**

**The search for greener ethylene**

Meanwhile, IBM **launched** a self-driving—or autonomous—lab combining AI with robotics at its research facility in Zurich. And Eli Lilly and Company rang in 2020 with the debut of a self-driving lab at its biotechnology center in San Diego.

"My lab has already been able to close the loop," Aspuru-Guzik says, describing a circuit of continuous learning in which AI algorithms guide data analysis and automation toward the identification, synthesis, and validation of novel molecules. The autonomous lab accesses, produces, and reprocesses data as it goes along.

"You ask me why I'm doing this; it's because the world has no time.

— ***Alán Aspuru-Guzik,*** *professor of chemistry and computer science, University of Toronto*

These labs aren't perfect yet, and much work and convincing are needed before research managers in the drug and chemical industries embrace them. Observations vary on where the bottlenecks lie in the clocklike loops—some point to the data; others, to the robots. But the innovators are unanimous regarding the importance of the third element in the loop: the human researcher.

"This idea of the clockwork laboratory is, in the long term, not the strongest approach," Cooper says of the notion of a self-sufficient research machine. "The strongest approach is to have the clockwork laboratory with a very permeable interface so that the human knowledge can be captured as well."

## THE ACADEMICS

**AI** most likely debuted in science fiction with Samuel Butler's 1872 novel *Erehwon, or Over the Range*. It made news in the real world a quarter century ago when IBM's Deep Blue supercomputer **took down** Garry Kasparov, the world chess champion. Kasparov came to terms with his defeat, much as the research community has learned to stop worrying and love a machine that accelerates discovery.

"You ask me why I'm doing this; it's because the world has no time," Aspuru-Guzik says. He points to rapid design techniques in industries such as automotive that rely on advanced materials and the urgency to confront climate change with new materials for storing sun and wind energy. "We have to enter the era of rapid prototyping in materials."

Researchers say the challenge in accelerating discovery comes down to improving their data sets. That's a primary function of machine learning in an autonomous lab as it cycles and directs data from synthesis and validation, melding them with data from available published literature.

Assembling and fine-tuning a system in Toronto to demonstrate the power of closed-loop autonomous discovery took about a year and a half, Aspuru-Guzik says. Once operational, the machine was able to produce about 40 molecules in a production run, he says, "which is more than the number of published molecules in the field I am working on, organic semiconductors for laser devices."

Credit: Evonik Industries

A scientist at Evonik Industries prepares a high-throughput screening apparatus at the company's laboratory in Essen, Germany. Evonik is investigating the integration of automation and artificial intelligence.

But while the system is intended to support an around-the-clock process, it is still operating in discrete runs, each averaging a day and a half. “The biggest holdup is not AI,” Aspuru-Guzik says. “Data bottlenecks? Zero. The area that is the challenge for getting the self-driving lab to work is the synthesis machinery. The robotics are a little finicky, a little hard to control.”



Those robots are the source of the synthesis and validation data necessary for machine-learning algorithms to drive toward discovery. “The point is to generate data on the fly,” he says. “That is what you’re getting a robot for.”

The University of Liverpool applied a twist to robotics in its autonomous lab. “It’s ironic that Alán calls his a self-driven laboratory,” Cooper quips. “Our robots actually drive.” Indeed, robotic agents scoot around a traditional-looking research space in a **video** released by Cooper’s group. “We decided to automate the chemist, not the instrument,” he says.

Cooper’s catalyst experiment operated nonstop for 8 days, completing about 6,500 manipulations in a complex workflow involving solid and liquid handling, some under a nitrogen-sealed, inert atmosphere, and multiple measurements. The mobile robots scooted a total of 2.2 km in a room measuring 5 by 10 m. Cooper says the system can theoretically operate for much longer. “We used 8 different experimental stations,” he says. “It could have been 18 or 80.”

As for bottlenecks, Cooper also points to robot mechanics. “AI is most powerful when there are multiple choices, a wide variety of measurements,” he says. But the more involved the experiment, the more complex the machinery, “and every bit needs to be really reliable. Once you string 10 operations together, even a failure point of 0.1% becomes quite significant.”

Regina Barzilay, a computer science professor at the Massachusetts Institute of Technology and colead of the Machine Learning for Pharmaceutical Discovery and Synthesis Consortium, notes that AI can run into problems as it **operates in new areas of chemical space**. But the technology works to solve the problems, partly through interaction with a researcher.



“You need to have a mechanism whereby the machine can tell you it needs to look in a particular part of chemical space where it needs to have more training,” Barzilay says. “An essential component is to know when the machine is actually not confident in its own prediction, which means it needs more help.”

Connor W. Coley, assistant professor of chemical engineering at MIT and part of the consortium, says advances made in pursuit of the autonomous laboratory reveal the critical importance of the human element. Even if the lab is a closed loop, it intersects with an adjacent loop defined by researcher interaction with the machine, starting from the word “Go.”

## LOOPED INTELLIGENCE

**An autonomous chemistry laboratory runs experimental cycles intended to yield useful molecules. In the cycle, artificial intelligence models the experiment and designs a compound, robotic equipment runs the synthesis, and AI evaluates the output; researchers interpret the data and adjust experimental models or the goal definition as needed.**

**Credit: Adapted from Connor W. Coley/Will Ludwig/C&EN**

**Source:** Connor Coley/Massachusetts Institute of Technology

“Humans are always going to set the design objective and specify something that an algorithm can reduce to a numerical optimization,” Coley says. “Humans will always be setting the big-picture goal” through rounds of interrogation.

“The big questions include which experiments are accessible to the platform compared to which are needed to prove or disprove the kinds of hypotheses we’re after,” he continues. “Another big question is which are the workflows that we can physically automate as steps in the process.”

And there are considerations beyond academic labs like Coley’s. “The interesting question is whether industry is actually going to bite on this or whether it’s some academic bullshit,” Cooper says. His laboratory is exploring this question directly: Mobotics, a company spun off from Cooper’s lab, is pursuing business with materials companies.

Similarly, Aspuru-Guzik colaunched **Kebotix**, an AI services firm working with robotics, in 2017 when he was at Harvard University. He is also coordinating the launch of a consortium, called Acceleration Consortium, focused on autonomous technology in materials discovery. Members include industrial and academic researchers.

## THE START-UPS

IBM followed up its chess win with a second popular showcase for AI by **introducing Watson**, a question-answering computer, as a contestant on the quiz show *Jeopardy!* In 2011, Watson defeated champions Brad Rutter and Ken Jennings, winning $1 million.

IBM went on to launch Watson as a commercial product; its first application was in decision management for lung cancer treatment at Memorial Sloan Kettering Cancer Center. The firm also customized Watson for industrial research markets, including chemicals, in which companies including the big German firm Evonik Industries have deployed it.

0:00 / 0:30


Credit: IBM

The robotic functions of RoboRXN at IBM's lab in Zurich are directed by artificial intelligence. AI-control data and graphics in the column on the left in this video illustrate intelligent control of the automated processes.

In 2019, IBM announced that it would shift development in its Watson Health division from drug discovery, in which the tool struggled to gain traction, **to clinical development**. And the company is working on a new AI architecture to support autonomous chemical discovery in both materials and drug applications.

IBM's RoboRXN for Chemistry combines AI algorithms, commercially available robotics, and cloud computing. The firm debuted RoboRXN in a fully autonomous lab last year at its Zurich research laboratory with experiments involving photoacid generator molecules, carbon-capture materials, and pharmaceutical compounds.

IBM began working on the system in 2017 with a project to apply natural language programming to predictive chemical synthesis problems. Philippe Schwaller, a PhD student at IBM who studied reaction prediction at the Zurich lab, says the AI component of RoboRXN digitizes chemistry through language to promote machine learning.



"Our models are trained on molecular representations where atoms are like characters, molecules are like words, and chemical reactions are like sentences," Schwaller says. Thus programmed, the model can learn in a fashion that mimics human learning.

The model is also generating data, says Teodoro Laino, a research scientist and leader of the RoboRXN project. "AI is generating instructions," he says. "The net effect of AI is to add software that is writing software for the robot."

IBM is working with industrial and institutional partners to implement the technology in research, Laino says. For example, XChem, an experimental facility at Diamond Light Source, the UK's national synchrotron facility, is testing whether RoboRXN can discover compounds from data the group generates on how small molecules bind to proteins.

IBM has competition in offering AI research services: start-ups like Insilico Medicine, Exscientia, and Citrine Informatics are landing contracts with big chemical and pharmaceutical firms. Some of them are also teaming up with lab robotics providers.

Insilico, for example, announced a partnership last July with Arctoris, a supplier of automated drug discovery technology. Insilico CEO Alex Zhavoronkov says the company is looking for an optimum fit between AI and robotics in a closed-loop laboratory. "We are trying to get high-quality data from a controlled robotics environment," he says.

For Zhavoronkov, as for Aspuru-Guzik at the University of Toronto, robots are the crux. "Synthesis is the main bottleneck. Currently we outsource synthesis" to a contract research organization (CRO), he says. Insilico used AI to discover an **idiopathic pulmonary fibrosis drug candidate** earlier this year, then signed on the CRO WuXi AppTec to synthesize it for testing.

> "**'Self-guided' in this area means within boundaries set by human scientists.**
>
> — ***Paul Collier,*** *research fellow, Johnson Matthey*

Exscientia launched in 2012 with a vision of automated drug design, CEO Andrew Hopkins says. The firm is continually searching for elements of the process in which AI can increase productivity, he says.

"Drug discovery is both a big data and a small data problem," Hopkins says. Databases, patents, and pertinent literature create a huge amount of data from which to build models, he says. "But whenever we look at first-in-class drug targets, chances are we know very little about them." AI allows researchers to design algorithms that explore chemical space beyond the parameters of the original experimental model. "We think of drug discovery as a learning problem, not a screening problem," he says.

The company, which counts **GlaxoSmithKline as a client**, is preparing to bring robots into the fold. “We are building labs in this space and hiring a director of automation,” Hopkins says.

Citrine targets the materials industry and boasts **the chemical giant BASF as a client**. The company is keeping its focus on AI, joining where needed with partners for robotics and other components of a closed-loop discovery system. CEO Greg Mulholland describes the role of AI in an autonomous lab as a central guidance system working hand in hand with the researcher. Other elements of the loop will vary from company to company, as will the closedness of the loop, Mulholland says.

## THE REAL-WORLD USERS

“Artificial intelligence is very much on our agenda,” says Henrik Hahn, Evonik’s chief digital officer. The German chemical company has been working for several years to understand the implications of AI in materials discovery as part of a **digitization** program that has introduced AI at various stages of research. The firm is moving toward closed-loop AI and robotics, but Hahn questions whether such an environment will ever fully evolve.

“The autonomous laboratory in materials discovery appears to be some kind of holy grail as computer programs and algorithms surpass the creativity of our materials science experts,” he says. “But this is really rather a vision, and a vision means it will never come true.”


**Credit: Exscientia**

Patrick Collins, lab operations manager at Exscientia employs artificial intelligence to search for potential drug compounds. CEO Andrew Hopkins describes drug discovery as “a learning problem, not a screening problem.”

Evonik is using IBM’s Watson AI technology in its labs and programing its own algorithms, Hahn says. And it is beginning to explore automation, looking, for example, into IBM’s RoboRXN technology.

Hahn emphasizes that the autonomous lab poses a steep management challenge. “Wherever we can automate, we will try to do so,” he says, though costs can be an issue. And the people who work in the lab now are no small consideration. Lab automation “is clearly disrupting classic lab work,” Hahn says. “This term *autonomous lab* has to be handled with care because there will always be the connotation that we are substituting human beings.”

Dow is similarly pacing itself toward achieving autonomy in the lab. “My vision of the lab of the future has been practiced at Dow for a while,” says A. N. Sreeram, the firm’s chief technology officer. He notes that Dow built and programmed its own supercomputers since late in the first decade of this century and is among the industry pioneers in automated research.



These days, Sreeram adds, the firm is working with robotics, big data, high-performance computing, and now AI and machine learning. And the company has been accessing technology through partnerships. They include a 2017 pact with 1QBit to develop quantum computing applications and a deal announced last year with Microsoft to develop AI for polyurethane research.

The British chemical maker Johnson Matthey is working with Mobotics, the company launched by Cooper at the University of Liverpool, on robotics for data-driven discovery. The firm says the work builds on a tradition of enabling scientists in materials discovery through technology that supports synthesis, characterization, and measurement.



“The self-driving or self-guided laboratory is another practical capability which our scientists will use to speed up the innovation process,” Paul Collier, a research fellow at the company, says in an email. “ ‘Self-guided’ in this area means within boundaries set by human scientists.” AI-directed automation provides a tool “that complements human scientists rather than replacing them,” he says.

Another chemical maker, DSM, recently announced a partnership with Delft University of Technology aimed at linking AI with automation. Marcus Remmers, DSM’s chief technology officer, says the project aims to supercharge aspects of industrial biotech research—automation, modeling, data management, and AI—that are already in place.

> **“This term *autonomous lab* has to be handled with care because there will always be the connotation that we are substituting human beings.**
>
> — ***Henrik Hahn,*** *chief digital officer, Evonik Industries*

A major focus will be on the researcher, Remmers says. “In the early stages of our digital transformation in R&D, we were focused on pilots and tools, trying to work things out in small environments,” he says. “As we mature, we realize the tools are not the limiting factor any more. More so, it is the mindset of the people. The scientist will have to reinvent what they are and how they see themselves and the value that they bring to this new world. If you see yourself as just somebody setting up a machine, you may well end up missing the big picture.”

In drug discovery, AI has gained traction managing data complexity. “Huge progress has been made over the last few years in comprehensively cleaning, unlocking, and harnessing the diverse and large volumes of discovery data accumulated over decades of research,” Hugo Ceulemans, scientific director of discovery data sciences at Janssen Research & Development, says in an email.

Drug companies are investing in data generation and acquisition for traditional data types such as assays and chemical reactions, as well as for new types, such as high-content microscopy images. As a result, Ceulemans says, pharmaceutical research is building up an AI and automation infrastructure.

“The large data volumes flowing from automated pipelines boost AI impact on the portfolio,” he says. AI-created insights are also beginning to direct the data that scientists collect, he adds.

Biotech companies have the luxury of implementing AI from the ground up. Moderna’s success in arriving at a vaccine with 95% efficacy against COVID-19 in less than a year is partly attributed to AI algorithms, according to an **article** in Digital Initiative, published by Harvard Business School.

For established drug companies, the adoption of AI can be a more protracted journey. Some major companies, however, are making a determined effort to close the research technology loop.

Eli Lilly and Company opened a robotics center in San Diego early last year in partnership with Strateos, a developer of research-scheduling software. The Lilly Life Sciences Studio is a 1,100 $m^2$ facility that includes an autonomous lab with over 100 instruments and storage of over 5 million compounds. It is the culmination of a 6-year project, says James P. Beck, the head of medicinal chemistry at the center.



Like most drug companies, Lilly has been doing automated biology for decades and automated chemistry for more than 10 years, Beck says. The Life Sciences Studio puts the company’s expertise in chemistry, in vitro biology, sample management, and analytical data acquisition in a closed loop. AI controls robots that Lilly researchers can access via the cloud, he says.

The lab, which is operated by Strateos, is also accessible to outside researchers, who can bring their own data, compounds, and experiments to the system.

Christos A. Nicolaou, head of chemical informatics at Lilly, says AI algorithms have evolved to the level where they can orchestrate automated operations in a lab. “AI is mature enough nowadays, and we have enough good data to come down to earth and design with action in mind,” he says. Lilly worked with software developers, Nicolaou says, but designed the AI architecture in-house.

“The lab of the future is here today,” Beck says. But closing the loop requires heavy lifting. “It is a multifactorial challenge involving science, hardware, software, and engineering,” he says. “It is far more than a science story.”

In fact, proponents of the autonomous lab suggest it is a human evolution story—one in which a technological environment rises around the enlightened scientist, posing little threat to the human. “Imagination and creativity will remain human for the foreseeable future,” the University of Liverpool’s Cooper says.

Aspuru-Guzik at the University of Toronto takes that idea one step further, quoting Jorge Luis Borges’s poem “Chess”:

*God moves the player as he the pieces*
*But what god behind God plots the advent*
*Of dust and time and dreams and agonies?*

That god, Aspuru-Guzik says, is human.

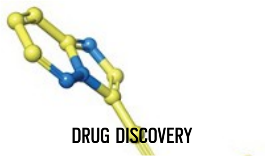
Chemical & Engineering News
ISSN 0009-2347

# YOU MIGHT ALSO LIKE...


**IBM debuts chemical synthesis robot**

INFORMATICS

**University of Leeds lab taps DeepMatter for lab automation**




**AI for drug discovery came into its own**


**ACS eBooks evidence-based acquisition**
by ACS Publications

# JOIN THE CONVERSATION

**Contact the reporter**

**Submit a Letter to the Editor for publication**

**Engage with us on Twitter**