# Artificially intelligent 'Coscientist' automates scientific discovery

December 20 2023



The Carnegie Mellon University Cloud Lab is a remotely operated, automated lab that gives researchers access to more than 200 pieces of scientific equipment. Credit: Carnegie Mellon University

A non-organic intelligent system has for the first time designed, planned

and executed a chemistry experiment, Carnegie Mellon University researchers [report](link) in the Dec. 21 issue of the journal *Nature*.

"We anticipate that intelligent agent systems for autonomous scientific experimentation will bring tremendous discoveries, unforeseen therapies and new materials. While we cannot predict what those discoveries will be, we hope to see a new way of conducting research given by the synergetic partnership between humans and machines," the Carnegie Mellon research team wrote in their paper.

The system, called Coscientist, was designed by Assistant Professor of Chemistry and Chemical Engineering Gabe Gomes and chemical engineering doctoral students Daniil Boiko and Robert MacKnight. It uses large language models (LLMs), including OpenAI's GPT-4 and Anthropic's Claude, to execute the full range of the experimental process with a simple, plain language prompt.

For example, a scientist could ask Coscientist to find a compound with given properties. The system scours the Internet, documentation data and other available sources, synthesizes the information and selects a course of experimentation that uses robotic application programming interfaces (APIs). The experimental plan is then sent to and completed by automated instruments. In all, a human working with the system can design and run an experiment much more quickly, accurately and efficiently than a human alone.

"Beyond the chemical synthesis tasks demonstrated by their system, Gomes and his team have successfully synthesized a sort of hyper-efficient lab partner," says National Science Foundation (NSF) Chemistry Division Director David Berkowitz. "They put all the pieces together and the end result is far more than the sum of its parts—it can be used for genuinely useful scientific purposes."

Specifically, in the *Nature* paper, the research group demonstrated that Coscientist can plan the [chemical synthesis](#) of known compounds; search and navigate hardware documentation; use documentation to execute high-level commands in an automated lab called a cloud lab; control liquid handling instruments; complete scientific tasks that require the use of multiple hardware modules and diverse data sources; and solve optimization problems by analyzing previously collected data.

"Using LLMs will help us overcome one of the most significant barriers for using automated labs: the ability to code," said Gomes. "If a scientist can interact with automated platforms in natural language, we open the field to many more people."



The Carnegie Mellon University Cloud Lab is a remotely operated, automated lab that gives researchers access to more than 200 pieces of scientific equipment.

This includes [academic researchers](#) who don't have access to the advanced scientific research instrumentation typically only found at top-tier universities and institutions. A remote-controlled automated lab, often called a cloud lab or self-driving lab, brings access to these scientists, democratizing science.

The Carnegie Mellon researchers partnered with Ben Kline from [Emerald Cloud Lab](#) (ECL), a Carnegie Mellon-alumni founded, remotely operated research facility that handles all aspects of daily lab work, to demonstrate that Coscientist can be used to execute experiments in an automated robotic lab.

"Professor Gomes and his team's ground-breaking work here has not only demonstrated the value of self-driving experimentation, but also pioneered a novel means of sharing the fruits of that work with the broader scientific community using cloud lab technology," said Brian Frezza, co-founder and co-CEO of ECL.

Carnegie Mellon, in partnership with ECL, will open the first cloud lab at a university in early 2024. The Carnegie Mellon University Cloud Lab will give the university's researchers and their collaborators access to more than 200 pieces of equipment. Gomes plans to continue to develop the technologies described in the Nature paper to be used with the Carnegie Mellon Cloud Lab, and other self-driving labs, in the future.

Coscientist also, in effect, opens the "black box" of experimentation. The system follows and documents each step of the research, making the work fully traceable and reproducible.

"This work shows how two emerging tools in chemistry—AI and automation—can be integrated into an even more powerful tool," says Kathy Covert, director of the Centers for Chemical Innovation program at the U.S. National Science Foundation, which supported this work.

"Systems like Coscientist will enable new approaches to rapidly improve how we synthesize new chemicals, and the datasets generated with those systems will be reliable, replicable, reproducible and re-usable by other chemists, magnifying their impact."

Safety concerns surrounding LLMs, especially in relation to scientific experimentation are paramount to Gomes. In the paper's supporting information, Gomes's team investigated the possibility that the AI could be coerced into making hazardous chemicals or controlled substances.

"I believe the positive things that AI-enabled science can do far outweigh the negatives. But we have a responsibility to acknowledge what could go wrong and provide solutions and fail-safes," said Gomes.

"By ensuring ethical and responsible use of these powerful tools, we can continue to explore the vast potential of large language models in advancing scientific research while mitigating the risks associated with their misuse," the authors wrote in the paper.

Provided by Carnegie Mellon University