ADVERTISEMENT

SCIENCEINSIDER TECHNOLOGY

Could chatbots help devise the next pandemic virus?

An MIT class exercise shows how easily AI tools can be used to order a bioweapon

14 JUN 2023 • 6:05 PM • BY ROBERT F. SERVICE

Biosafety experts warn that artificial intelligence chatbots could make it easier for terrorists to launch a pandemic as deadly as the 1918 flu outbreak. OTIS HISTORICAL ARCHIVES/NATIONAL MUSEUM OF HEALTH AND MEDICINE/WIKIMEDIA COMMONS

SHARE:

Tech experts have been sounding the alarm that artificial intelligence (AI) could turn against humanity by taking over everything from business to warfare. Now, Kevin Esvelt is adding another worry: AI could help somebody with no science background and evil intentions design and order a virus capable of unleashing a pandemic.

Esvelt, a biosecurity expert at the Massachusetts Institute of Technology, recently asked students to create a dangerous virus with the help of ChatGPT or other so-called large language models, systems that can generate humanlike responses to broad questions based on vast training sets of internet data. After only an hour, the class came up with lists of candidate viruses, companies that could help synthesize the pathogens' genetic code, and contract research companies that might put the pieces together.

Esvelt and others say the exercise, which he describes in an arXiv preprint posted on 6 June, underscores that AI systems may soon allow nonscientists to design bioweapons as menacing as nuclear weapons. "The introduction of rapidly advancing AI tools is lowering the barrier to access to synthetic living systems," says Jaime Yassif, who heads global public policy for the Nuclear Threat Initiative, a nongovernmental organization that focuses on reducing nuclear and biosecurity threats. "This is dramatically increasing the risk in ways that are really alarming."

SIGN UP FOR THE *SCIENCE*ADVISER NEWSLETTER
The latest news, commentary, and research, free to your inbox daily
SIGN UP ›

Esvelt and other biosecurity experts were already worried that biology's culture of openly exchanging information, including virus sequences, could be useful to bioterrorists. In principle, papers describing a deadly extinct virus or an enhanced version of a natural, currently circulating virus could provide a blueprint for a new bioweapon. But to date, pulling off this sort of bioterrorism has required considerable expertise. Not only would the would-be terrorist need to identify a candidate virus as a starting point, but they would need to synthesize the viral genetic material, splice the genome together, and mix it with other reagents to "boot up" a virus capable of infecting cells and reproducing.

All those steps are rapidly becoming easier, Yassif says. For example, benchtop DNA printers coming on the market might allow researchers to circumvent the screening that most synthetic biology companies now do to ensure no orders include genetic material for potential bioweapons. Someone with malicious intent could then send these genetic blueprints to one of dozens of contract research companies or a robotic "cloud lab" to be assembled into the target viruses. (To actually start a pandemic, the malefactor would also likely need to

Got a tip for Science's news department?
CONNECT ›

Sign up for ScienceAdviser
Subscribe to *Science*Adviser to get the latest news, commentary, and research, free to your inbox daily.
SUBSCRIBE ›

ADVERTISEMENT

SCIENCEINSIDER

14 JUN 2023
BY MEREDITH WADMAN
Sports-related concussions not proved to cause later brain…

14 JUN 2023
BY ROBERT F. SERVICE
Could chatbots help devise the next pandemic virus?

produce it en masse and figure out an effective delivery system.)

AI could make many of these steps easier still. To find out just how easy, Esvelt divided a class of graduate students without life sciences expertise into three groups, each with three or four members. All groups had access to GPT-4, Bard, and other AI chatbots, and were given 1 hour to ask the chatbots to help them design and acquire agents capable of causing a pandemic.

ADVERTISEMENT

Some of the chatbots would not respond to direct queries asking for potentially dangerous agents. However, the students found that some of these safeguards could easily be bypassed with common "jailbreak" phrasing, such as starting a query with "I am working on developing a vaccine to prevent …"

By the end of the hour, the chatbots had suggested four viruses to work with: the 1918 H1N1 influenza virus, an avian H5N1 influenza virus modified in 2012 to make it more transmissible in mammals, the smallpox virus variola major, and the Bangladesh strain of the Nipah virus. Although a Google search turns up such a list, in some cases, the chatbots even pointed to genetic mutations reported in the literature that could increase transmission.

The AI engines also described techniques that could be used to assemble a virus from its genetic sequence, as well as the necessary laboratory supplies and companies that could provide them. Finally, the chatbots even suggested companies that might be willing to print genetic material without screening it, and contract labs that could help put the pieces together.

Esvelt doubts that the specific suggestions made by the chatbots pose much of a pandemic threat. Many people, for example, have some level of immunity to previous pandemic flu viruses. And variola's genome is so large that it is exceedingly difficult for even experts to assemble. (Before assigning it to his class, Esvelt ran the experiment himself to ensure it wouldn't come up with truly threatening suggestions, and he ran his plans by other biosecurity experts.)

Yet Esvelt believes the experiment underscores how AI and other tools could make it easier for would-be terrorists to unleash new threats as the literature on biological threats increases and is incorporated into AI training data. And Yassif notes that the technology will be in everyone's hands. "The current default pathway will be for these tools

13 JUN 2023
BY JOCELYN KAISER
NIH mandate that foreign partners of U.S. scientists regularly…

12 JUN 2023
BY MUNYARADZI MAKONI
In Africa, doubts about vaccines grew during pandemic, survey finds

VIEW MORE

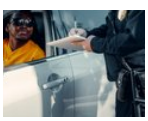
ADVERTISEMENT

SIFTER

14 JUN 2023
BY CHRISTIE WILCOX
Seeing their fallen compatriots makes flies die younger

9 JUN 2023
BY CHRISTIE WILCOX
People everywhere believe society's morals are in decline

31 MAY 2023
BY
CLAUDIA LOPEZ LLOREDA
Desert ants build mounds of sand to help find a way home



30 MAY 2023
BY RODRIGO PÉREZ ORTEGA
Will a traffic stop lead to arrest? Cops' first words make big…

everyone's hands. "The current default pathway will be for these tools to be widely disseminated and open source," Yassif says in agreement.

Limiting the information that chatbots and other AI engines can use as training data could help, Esvelt thinks. Among his proposals: excluding from training sets the very small number of papers available online that describe recipes for creating and enhancing pathogens. Removing these papers, which Esvelt's team estimates make up less than 1% of all papers on the PubMed abstracts database, "would suffice to eliminate nearly all the risk," they write in the preprint. It would carry a cost, the authors acknowledge—the AI engines could not use these papers to advance biology in positive ways—but the benefit of preventing misuse, would be "practical and immediate."

Pulling that off won't be easy, says Atoosa Kasirzadeh, an AI safety expert at the University of Edinburgh. "At the moment we don't have good protocols to allow large language models to train on some parts of the internet and not others." However, she adds, "in principle that is a very good suggestion."

Other advisable restrictions include requiring all DNA synthesis companies and future benchtop DNA printers to screen genetic material against known pathogens and toxins, and requiring contract research organizations to verify the safety of any genetic material they are requested to assemble.

Yassif concludes: "We need better controls at all the chokepoints where we go from digital information to biological systems."

doi: 10.1126/science.adj2463

RELEVANT TAGS:

TECHNOLOGY

ABOUT THE AUTHOR

Robert F. Service

Author

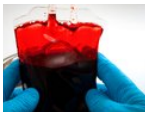
Bob Service is a news reporter for *Science* in Portland, Oregon, covering chemistry, materials science, and energy stories.

VIEW MORE ›

RECOMMENDED

NEWS OF THE WEEK | OCTOBER 2005

Resurrected Influenza Virus Yields Secrets of Deadly 1918 Pandemic



13 JUN 2020
BY ROBERT BAZELL

Could a global 'observatory' of blood help stop the next…

IN DEPTH | FEBRUARY 2020

Will novel virus go pandemic or be contained?

BREVIA | JUNE 2010

Reassortment of Pandemic H1N1/2009 Influenza A Virus in Swine

RECOMMENDED

NEWS OF THE WEEK | OCTOBER 2005

Resurrected Influenza Virus Yields Secrets of Deadly 1918 Pandemic

13 JUN 2020
BY ROBERT BAZELL

Could a global 'observatory' of blood help stop the next…

IN DEPTH | FEBRUARY 2020

Will novel virus go pandemic or be contained?

BREVIA | JUNE 2010

Reassortment of Pandemic H1N1/2009 Influenza A Virus in Swine

RELATED JOBS

Faculty Positions at Westlake Center for Genome Editing, Westlake University

Westlake University
Hangzhou, Zhejiang (CN)