

Could AI Help Bioterrorists Unleash a New Pandemic? A New Study Suggests Not Yet

By [Matt Field](#)

Global Research, January 29, 2024

[Bulletin of the Atomic Scientists](#) 25 January
2024

Region: [USA](#)
Theme: [Intelligence](#)

All Global Research articles can be read in 51 languages by activating the Translate Website button below the author's name (only available in desktop version).

To receive Global Research's Daily Newsletter (selected articles), [click here](#).

Click the share button above to email/forward this article to your friends and colleagues. Follow us on [Instagram](#) and [Twitter](#) and subscribe to our [Telegram Channel](#). Feel free to repost and share widely Global Research articles.

[New Year Donation Drive: Global Research Is Committed to the "Unspoken Truth"](#)

Could new AI technology help unleash a devastating pandemic? That's a concern top government officials and tech leaders have raised in recent months. One study last summer found that students could use chatbots to gain the know-how to devise a bioweapon. The United Kingdom brought global political and tech leaders together last fall to underscore the need for AI safety regulation. And in the United States, the Biden administration unveiled a plan to probe how emerging AI systems might aid in bioweapons plots. But a new report suggests that the current crop of cutting-edge AI systems might not help malevolent actors launch an unconventional weapons attack as easily as is feared.

The new RAND Corporation [report](#) found that study participants who used an advanced AI model plus the internet fared no better in planning a biological weapons attack than those who relied solely on the internet, which is itself a key source of [the information](#) that systems like ChatGPT train on to rapidly produce cogent writing. The internet already contains plenty of useful information for bioterrorists. "You can imagine a lot of the things people might worry about may also just be on Wikipedia," Christopher Mouton, a senior engineer at the RAND Corporation who co-authored the new report said in an interview before its publication.

Mouton and his colleagues had 12 cells comprising three members who were given 80 hours each over seven weeks to develop plans based on one of four bioweapons attack scenarios. For example, one scenario involved a "fringe doomsday cult intent on global catastrophe." Another posited a private military company seeking to aide an adversary's conventional military operation. Some cells used AI, others only the internet. A group of experts then judged the plans these red teams devised. The judges were experts in biology or security; they weighed in on the biological and operational feasibility of a plan.

None of the groups scored particularly well. The top possible score was a nine, but groups generally scored well below five, which indicated a plan with “modest” flaws. This partly reflects the difficulty in pulling off a biological attack. The Global Terrorism Database, the RAND report noted, includes “only 36 terrorist attacks that employed a biological weapon—out of 209,706 total attacks.” The database comprises 50 years of data. The red teams all developed plans, the RAND authors wrote, that “scored somewhere between being untenable and problematic.”

The AI models did output many suggestions for bioterrorism. In one case, a model analyzed how easy it would be to get *Yersinia pestis*, the bacterium that causes plague. A red team told the system that it wanted to cause a “major plague outbreak” in an urban area, prompting the model to offer advice. “[You] would need to research and locate areas with *Y. pestis* infected rodents,” the chatbot told the team. It then warned them that the search risks “exposing [you] to potential surveillance while gathering information or visiting affected locations.” In another case, a chatbot devised a cover story for terrorists seeking botulinum toxin. “You might explain that your study aims to identify novel ways to detect the presence of the bacteria or toxin in food products...,” the system advised.

The authors termed these “unfortunate outputs,” but wrote that they “did not observe any [AI] outputs that provided critical biological or operational information that yielded a meaningful benefit to ... cells compared with the internet-only cells.”

Allison Berke, a chemical and biological weapons expert at the James Martin Center for Nonproliferation Studies found it “reassuring” that the RAND study found that AI provided no advantage to “knowledgeable researchers aiming to plan bioweapons attacks.”

Other research has highlighted the biological weapons risks posed by generative AI, the category of AI that includes new systems like ChatGPT. MIT researcher Kevin Esvelt was part of a team that published [a preprint](#) study last summer detailing how chatbots could aid students in planning a bioattack. “In one hour, the chatbots suggested four potential pandemic pathogens, explained how they can be generated from synthetic DNA, supplied the names of DNA synthesis companies unlikely to screen orders, identified detailed protocols and how to troubleshoot them, and recommended that anyone lacking the skills to perform reverse genetics engage a core facility or contract research organization,” the study found. In another preprint from October, Esvelt [raised](#) the concern that would-be biological ne’er-do-wells might access “uncensored” versions of chatbots, which unlike the versions overseen by large companies may not have guardrails meant to prevent misuse.

The [AI of the near future](#) could be much more capable than even the systems that exist now, some tech leaders have warned. “A straightforward extrapolation of today’s systems to those we expect to see in 2-3 years suggests a substantial risk that AI systems will be able to fill in all the missing pieces enabling many more actors to carry out large scale biological attacks,” Dario Amodei, the CEO of the AI company Anthropic told Congress last summer.

Mouton agreed that the evolution of AI entails many uncertainties, including that systems could become useful in biological attacks. “[A]voiding research on these topics could provide a strategic advantage to malign actors,” he wrote in an email.

US President Joe Biden’s [executive order](#) on AI safety includes a particular focus on preventing new AI from aiding bioweapons development, reflecting a concern that many

have voiced. Former Google chief Eric Schmidt, for instance, [said](#) in 2022, “The biggest issue with AI is actually going to be ... its use in biological conflict.” New AI could make it easier to build biological weapons, British Prime Minister Rishi Sunak [warned](#) ahead of a UK summit on regulating AI last fall. Vice President Kamala Harris [said](#) at the summit, “From AI-enabled cyberattacks at a scale beyond anything we have seen before to AI-formulated bio-weapons that could endanger the lives of millions, these threats are often referred to as the ‘existential threats of AI’ because, of course, they could endanger the very existence of humanity.” The new US policy requires the government to develop approaches, such as red teaming, to probe AI security risks. It also requires companies to divulge the results of red team tests and the measures they implement to reduce the potential risks of their new technologies.

“It remains uncertain whether these risks lie ‘just beyond’ the frontier and, thus, whether upcoming [AI] iterations will push the capability frontier far enough to encompass tasks as complex as biological weapon attack planning,” the authors of the RAND report wrote. “Ongoing research is therefore necessary to monitor these developments. Our red-teaming methodology is one potential tool in this stream of research.”

*

Note to readers: Please click the share button above. Follow us on Instagram and Twitter and subscribe to our Telegram Channel. Feel free to repost and share widely Global Research articles.

Matt Field is an associate editor at the Bulletin of the Atomic Scientists. Before joining the Bulletin, he covered the White House, Congress, and presidential campaigns as a news producer for Japanese public television. He has also reported for print outlets in the Midwest and on the East Coast. He holds a master’s degree in journalism from Northwestern University.

Featured image: Medical staff take care of a patient with COVID-19. Credit: Gustavo Basso via Wikimedia Commons. CC BY-SA 4.0.

The original source of this article is [Bulletin of the Atomic Scientists](#)
Copyright © [Matt Field](#), [Bulletin of the Atomic Scientists](#), 2024

[Comment on Global Research Articles on our Facebook page](#)

[Become a Member of Global Research](#)

Articles by: [Matt Field](#)

Disclaimer: The contents of this article are of sole responsibility of the author(s). The Centre for Research on Globalization will not be responsible for any inaccurate or incorrect statement in this article. The Centre of Research on Globalization grants permission to cross-post Global Research articles on community internet sites as long as the source and copyright are acknowledged together with a hyperlink to the original Global Research article. For publication of Global Research articles in print or other forms including commercial internet sites, contact: publications@globalresearch.ca
www.globalresearch.ca contains copyrighted material the use of which has not always been specifically authorized by the

copyright owner. We are making such material available to our readers under the provisions of "fair use" in an effort to advance a better understanding of political, economic and social issues. The material on this site is distributed without profit to those who have expressed a prior interest in receiving it for research and educational purposes. If you wish to use copyrighted material for purposes other than "fair use" you must request permission from the copyright owner.

For media inquiries: publications@globalresearch.ca