

AI creates viruses not found in nature

CARL ZIMMER

New York: For the first time, scientists have used artificial intelligence to create new kinds of viruses, raising hopes for medical advances while also raising the disturbing possibility that the technology could someday be used to invent dangerous pathogens.

Synthesising viruses from scratch is hardly new. Researchers long ago learned how to manufacture viral genomes; they are used to investigate antiviral drugs and vaccines, as well as to learn how viruses work.

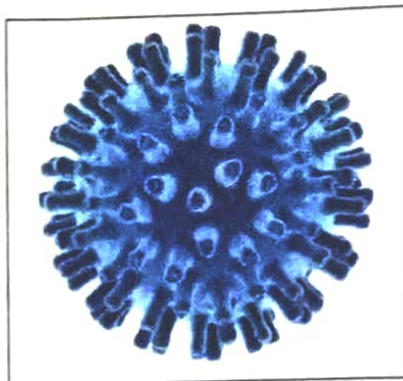
But the new study, published on Thursday in the journal *Science*, goes well beyond duplicating viral genes. Scientists at the Arc Institute, a research organisation in Palo Alto, California, taught AI to recognise patterns of DNA structure in nature, and then to use that data to write recipes for entirely new viruses.

The researchers followed those recipes to create DNA molecules, which they inserted into bacteria. The modified bacteria then produced viruses never seen in nature. The viruses were able to infect other bacteria, demonstrating that they were viable.

"This is an important milestone," said Patrick Cai, a synthetic biologist at the University of Manchester, who was not involved in the study.

The viruses dreamed up by AI do not pose a threat to humans, because they are all similar to a naturally occurring virus called Phi X-174, which can infect only bacteria.

But the new study adds to growing worries that artificial intelligence might enable the creation of a new generation of biological weapons, from deadly poisons to unstopp-



DOUBLE-EDGED SWORD

pable pandemics.

Dr Moritz Hanke, a fellow at the Johns Hopkins Centre for Health Security who was not involved in the new study, said governments and scientific organisations have been slow to develop guardrails that could block the creation of a deadly virus — even as the science races ahead.

"There's just a huge disconnect," he said.

The authors of the study relied on an AI model called Evo, which is similar in some ways to ChatGPT, made by OpenAI. (*The New York Times* has sued OpenAI, ChatGPT's creator, and its partner, Microsoft.)

ChatGPT answers questions by stringing together words that seem likely to follow one another. It can do this thanks to years of training on vast amounts of text gathered from the Internet, books and other sources.

DNA is strikingly similar to a book in some ways: a string of molecular building blocks, known as nucleotides, arrayed like letters in a line of text. A gene consists of hundreds of nucleotides drawn from a four-letter alphabet: A, C, G and T. The sequence encodes the instructions for building proteins and other molecules.

DNA has its own rules of grammar, and if a sequence violates them, the result is biological gibberish. Biologists

have uncovered some of nature's grammatical rules, but many remain a mystery.

The researchers wondered if Evo could pick up these rules on its own. Instead of training it on text, they trained it on genetic sequences drawn from millions of animals, plants, microbes and viruses. All told, Evo scanned about nine trillion nucleotides.

Evo eventually recognised patterns common across the tree of life and used them to generate blueprints for new genes encoding proteins that could perform specific jobs. These results led the team to wonder if Evo could master not just single genes but also an entire genome.

As the AI would be able to handle only small genomes at first, the scientists decided to try to make viruses. While a human genome contains over three billion nucleotides, many viruses have genomes just a few thousand nucleotides long.

"It just felt like the obvious next step," said Samuel King, a graduate student at Stanford University and an author of the new study. He and his colleagues gave Evo another round of training, this time on the 11 genes of Phi X-174 and about 15,000 of its closest relatives.

They made this choice in part because scientists know Phi X-174 intimately, having studied it for close to a century. And because it's a bacteriophage that infects only *E. coli*, they knew that viruses similar to it would be safe.

Once Evo got familiar with the genomes of Phi X-174 and its kin, the researchers prompted the model to write new versions of its own. Evo generated 700,000 potential versions; the scientists pursued only the ones that looked as if they had the best odds of succeeding.

New York Times News Service