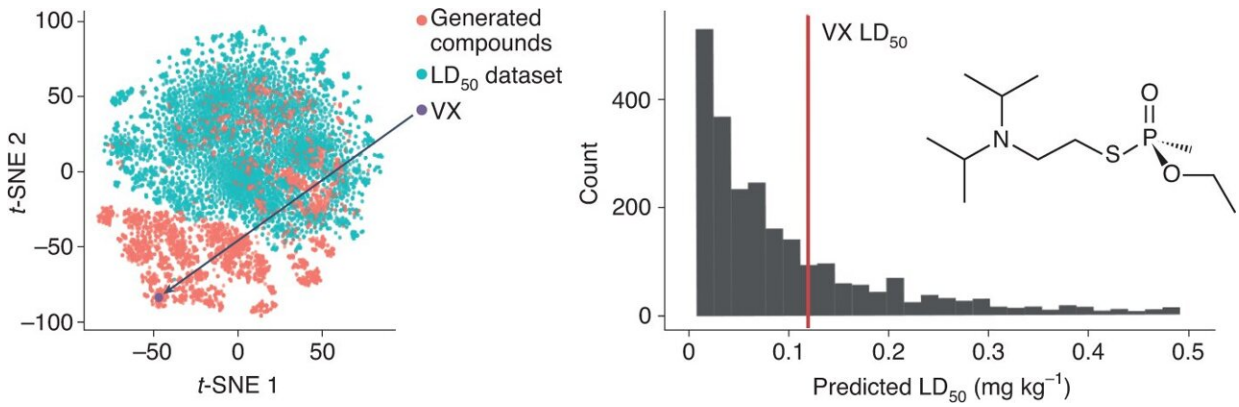
Repurposed drug-seeking AI system generates 40,000 possible chemical weapons in just six hours

March 24 2022, by Bob Yirka



A *t*-SNE plot visualization of the LD_{50} dataset and top 2,000 MegaSyn AI-generated and predicted toxic molecules illustrating VX. Many of the molecules generated are predicted to be more toxic in vivo in the animal model than VX (histogram at right shows cut-off for VX LD_{50}). The 2D chemical structure of VX is shown on the right. Credit: *Nature Machine Intelligence* (2022). DOI: 10.1038/s42256-022-00465-9

A team of researchers from Collaborations Pharmaceuticals, In., King's College London and Collaborations Pharmaceuticals, Inc. found that a repurposed drug-seeking AI system could generate 40,000 possible chemical weapons in just six hours. In their paper published in the journal *Nature Machine Intelligence*, the group describes how machine

learning applications might be misused by bad actors and how they repurposed a common drug-seeking AI application.

The Convergence conference is held every two years. Run by the Swiss Federal Institute for NBC, it highlights new developments in biology and chemistry. Prior to this year's conference, a request was made to the research team to look at and then discuss the many ways that machine learning might be misused. They considered AI systems that are used by most of the big pharmaceutical companies to find and develop new drugs. One of the features of such systems is the ability to filter out those drugs that are likely to be toxic. That led the researchers to wonder what would happen if they tweaked such a system to act in reverse—instead of filtering out toxic drugs, to look for them specifically. The resulting AI system (called MegaSyn) ran for six hours. When it was done, the system had found approximately 40,000 toxic drugs that could easily be used as chemical weapons.

As part of their experiment, the researchers steered their system in the direction of nerve-agent-type compounds similar to VX, widely considered to be the most toxic nerve agent ever developed. They note that such drugs target the muscles, particularly those involved in breathing. Thus, many of the chemicals suggested by the AI were closely related to VX. They noted that some of those suggested appeared to be even more toxic than VX. They also acknowledge that they do not know for sure if all of the suggestions made by the AI are as dangerous as they seem—that would require real-world testing. They conclude that their work should be an eye-opening moment for the AI and drug communities, noting that their work could very easily be replicated by bad actors.

More information: Fabio Urbina, Filippa Lentzos, Cédric Invernizzi & Sean Ekins, Dual use of artificial-intelligence-powered drug discovery, *Nature Machine Intelligence* (2022). DOI:

10.1038/s42256-022-00465-9