taken any practical steps to turn it into a product?
HOSSEINI: At present, the idea has not yet entered the production stages. I have researched similar devices and am examining the features I intend to incorporate, but building a functional prototype requires technical resources, collaboration with specialists, and capital. When I first presented this idea, I was unable to register it in my own name because I had not yet turned 18, so it was registered in my father's name.
I presented the idea at international competitions and won a silver medal at an inventors' competition held in Switzerland, where I also received positive feedback. I had even been invited to attend an event in Switzerland, but the timing, my studies, and the circumstances of that period prevented me from attending the exhibition. Consequently, I lost the opportunity to establish direct contact with people who might have been interested in investing.

Besides the device for the visually impaired, do you have other ideas in the healthcare sector?
HOSSEINI: I have designed another idea called "MaterniGuard AI" to assist expectant mothers. My idea is to place a belt around the mother's abdomen that provides her with information about the fetus's condition.
Alongside this device, a mobile app would also be developed, with each mother having her own dedicated user dashboard. Information about the pregnancy would be recorded in this dashboard, and, if the necessary infrastructure were established, the physician could also access the required information. The goal is to enable the mother to obtain more information about her pregnancy and alleviate the family's concerns.

What prompted you to consider designing such a device?
HOSSEINI: This idea occurred to me in light of the experiences my mother recounted about her pregnancies. I considered how pregnancy can be accompanied by considerable anxiety for some families. Sometimes, a mother feels that something has changed in her abdomen or notices an unusual sign and becomes concerned that something may have happened to the fetus.
I wanted to develop an idea that could provide the mother with more information and help her manage her concerns. I presented this idea at international competitions as well and won a silver medal at a Swiss invention and innovation competition.
My other idea is to design a system for storing and accessing individuals' medical records. I considered that when someone migrates to another country, they may be unable to access their previous medical records or provide the necessary information to a new physician.
Under such circumstances, diagnosing an illness may become more difficult or treatment may be delayed, particularly when the individual cannot accurately explain their medical history, medications, and previous treatments.
Under my proposed design, each person would have a user dashboard containing information about their illnesses, prescriptions, prescribed medications, and treatment history. Individuals would be able to access their medical records in different countries and share essential information with their physicians.

What are your plans for your academic and professional future?
HOSSEINI: I would like to study computer science at the University of Tehran and expand my knowledge of programming and artificial intelligence. Afterward, I intend to work on my ideas and bring them to fruition. One of my major goals is to establish a company in the future dedicated to ideation, invention design, and manufacturing; a company that could evaluate new ideas, turn them into functional prototypes, and ultimately bring them into production. I would like to serve my country in this way. I believe that Iran has considerable talent, and that, if appropriate resources and opportunities were made available to these individuals, they could accomplish valuable work.

Biosecurity in age of AI



By Doug Irving
Corporate communications officer at RAND

INTERVIEW

Steph Guerra is an expert in AI, security, and biotechnology. The question that drives her work at RAND may be one of the most consequential for this moment in the AI era: How can we harness the full potential of AI for science and medicine — but also ensure it never helps anyone build a biological weapon?
Her career began in the lab. She's a molecular biologist by training, whose earliest research focused on preventing cancer. But she's been working on issues of AI safety since long before "large language model" was a household term. She now leads RAND's portfolio of research at the intersection of AI and biotechnology.
In a recent study, Guerra and a team of researchers tested so-called "open-weight" AI models. These models are made publicly available for anyone to download, run, and modify. That can make them more transparent and accessible than hosted models like Claude or ChatGPT. But it also creates concerns about how they could be misused.

What did you find in your study of open-weight models and biosecurity risks?
GUERRA: It's important to say, first, that these open-weight models are a central part of the AI ecosystem. They support open research; they allow organizations or people to deploy their own models without relying on third parties. But there are some potential downsides. Once they're released, their developers can no longer enforce safety controls, and the underlying model can be customized for beneficial or potentially harmful purposes.
We evaluated a few different open-weight models. We found that their capabilities are quite comparable to those of frontier models. But we also found that tampering with them to remove their safeguards was actually pretty trivial and low-cost. Obtaining a highly capable open-weight model, and then removing the safeguards that tell it to refuse certain questions, is within reach of many capable actors.

Does that mean someone could use these models to help plan a biological attack?
Just getting information out of a model is one small step on what would be a long pathway to carrying out an attack. A bad actor would still have to obtain the materials; design, build, and test the biological agent; weaponize it; and then release it. There are technical, operational, and motivational barriers along that entire pathway.
One of the real challenges we have is that it's difficult to assess how an AI model could help. Could it motivate someone to cause harm? Could it help with the technical aspects, using the appropriate tools to design and enhance an agent? We can all agree that people should not be able to use AI models to develop a bioweapon. But it's really hard to test the end-to-end risk of that. And so, we rely on these snapshots that our team and others build to understand where and how a bad actor could use AI.

What are some practical steps that AI developers can take to reduce the risk?

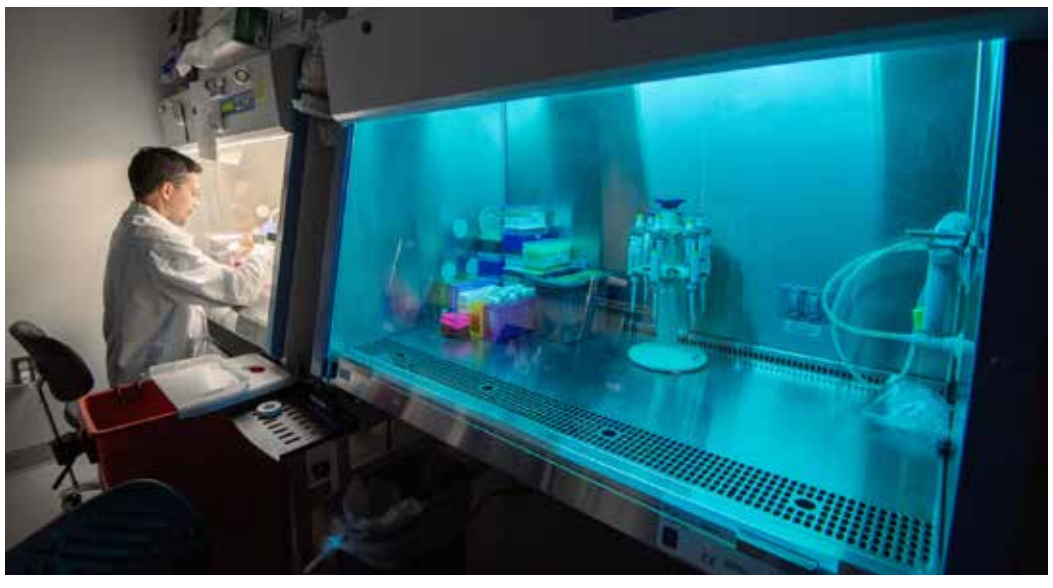


Steph Guerra

vietnam.vn

The ideal would be to build safeguards that are tamper-resistant. There has been some work on that, but it's unclear whether we'll ever be able to get there. If you think about it, closed-weight models are not tamper-free either. We always have to assume there will be ways to circumvent any single intervention put in place.

So then, what does an effective biosecurity strategy look like?
You need to create friction along the entire pathway. Our team looked at this problem and just recently developed a multi-layered strategy that would prevent a pandemic-scale biological attack.
It has four key theories of defense. One is managing access to information through things like stronger model safeguards and managed access programs. The second is preventing bad actors from getting access to the materials they would need to create a biological weapon, the DNA and RNA, such as through gene synthesis screening. The third is detecting misuse early through real-time monitoring. And the fourth is being able to rapidly respond to and attribute an attack, to deter a nefarious actor from attempting one in the first place.
Model safeguards might stop a lone actor, but by themselves, they won't slow down a sophisticated state with a bioweapons program. Early detection, early attribution — that deterrence layer — might. That's the challenge. There's never going to be a single silver-bullet solution that can stop all threats. We need all of these layers.



A scientist at Red Queen Bio grows cells to produce and test candidate antibody designs. The idea behind founding the startup company was to use AI to design remedies against AI-enabled bioweapons, and cure other scourges along the way, too.
WSJ

> We evaluated a few different open-weight models. We found that their capabilities are quite comparable to those of frontier models. But we also found that tampering with them to remove their safeguards was actually pretty trivial and low-cost. Obtaining a highly capable open-weight model, and then removing the safeguards that tell it to refuse certain questions, is within reach of many capable actors.

How do you avoid stifling legitimate research while still preventing malicious uses?
Biology is a dual-use endeavor. The same AI capabilities that might cause harm could also be incredibly useful for developing medical innovations. The goal here is managed access and not blanket prohibitions. It's to precisely define who a legitimate user is and then ensure those individuals have access to the most powerful tools and data they need.
There are some useful precedents. There are managed-access frameworks to address privacy and security concerns with human genome data, for example. It doesn't foreclose legitimate research, but it also doesn't put the data out there for anyone to misuse. Part of our philosophy at RAND is that the scientific community needs to be part of developing the solutions. AI safeguards are only going to work if they preserve meaningful utility for beneficial research.

Was there a moment in your career that put you on the path to doing what you do now?
A few years ago, I was part of a small federal team doing red-teaming experiments with AI models. I was watching these models answer questions about biological attacks, and they were providing really creative, useful information, fluently, and with a lot of confidence. And I remember it just struck me: We needed to be doing a lot more to monitor and track these risks.
I started my career as a cancer biologist. And I kept running into this reality that we pour years and resources into developing treatments, but not as much into preventing cancer in the first place. So, this idea of focusing on prevention, on getting ahead of risks so we're not scrambling to respond later, has just followed me through my career.

The interview was first published by RAND.