CSET CENTER for SECURITY and EMERGING TECHNOLOGY

AI and Biorisk: An Explainer

By Stephanie Batalis. PhD (sb2132@georgetown.edu)

Recent government directives, international conferences, and media headlines reflect growing concern that artificial intelligence could exacerbate biological threats.[1] When it comes to biorisk, AI tools are cited as enablers that lower information barriers, enhance novel biothreat design, or otherwise increase a malicious actor's capabilities.

It is important to evaluate AI's impact within the existing biorisk landscape to assess the relationship between AI-agnostic and AI-enhanced risks. While AI can alter the potential for biological misuse, focusing attention solely on AI may detract from existing, foundational biosecurity gaps that could be addressed with more comprehensive oversight.

Policies that effectively mitigate biorisks will also need to account for the varied risk landscape, because safeguards that work in one case are unlikely to be effective for all actors and scenarios. In this explainer, we outline the AI-agnostic and AI-enhanced biorisk landscape to inform targeted policies that mitigate real scenarios of risk without overly inhibiting AI's potential to accelerate cutting-edge biotechnology.

Our key takeaways regarding AI and biorisk include:

1. **Biorisk is already possible without AI, even for non-experts**. AI tools are not needed to access the foundational information and resources to cause biological harm. This highlights the need for layered safeguards throughout the process, from monitoring certain physical materials to bolstering biosafety and biosecurity training for researchers. The recent Executive Order on AI's requirement to screen DNA synthesis for federally-funded research is an example of a barrier to material acquisition.[2]
2. **The biorisk landscape is not uniform, and specific scenarios and actors should be assessed individually.** Distinct combinations of users and AI tools impact the potential for harm and the most effective likely policy solutions. Future strategies should identify clearly defined scenarios of concern and design policies to target them.
3. **Existing policies regarding biosecurity and biosafety oversight need to be clarified and strengthened.** AI-enabled biological designs are digital predictions that do not cause physical harm until they are produced in the real world. Such gain-of-function research, which modifies pathogens to be more dangerous, is already the target of existing policies.[3] However, these policies do

not adequately define what characteristics constitute risky research of concern, making them difficult to interpret and implement.[4] These policies are currently under review, and could be strengthened by establishing a standard framework of acceptable and unacceptable risk applicable to ***both*** AI-enhanced and AI-agnostic biological experimentation.[5]

These key takeaways are based on our analysis and summary of the biorisk landscape below.

Summary: AI's Impact on the Biorisk Landscape

Our assessment finds that scientific novices and experts alike can cause biological harm without access to AI. Adding AI tools changes the risk landscape in different ways depending on the specific scenario of misuse (Table 1).

Table 1. Impact of AI on Biological Misuse

	Without AI	With AI
Scientifically Naive User	- Finds and compiles information **themselves** using beginner-friendly online resources. - More likely to produce a **known, existing pathogen** or **toxin** than a new or modified one.	- A **chatbot** may help find and compile information. - More likely to produce a **known, existing pathogen** or **toxin** than a new or modified one. - Chatbots may lower the *perceived* barrier to information and engage new actors.
Scientifically Knowledgeable User	- Can follow existing scientific protocols to produce **known, existing pathogens** and **toxins**. - Can develop **novel or modified pathogens** or **toxins** through directed experimentation.	- Can use **biological design tools (BDTs)** to design **new or modified** pathogens and toxins, or to **evade screening** or **enhance production**. - BDTs may increase design efficiency and reduce the physical testing burden. - Chatbots may help to **brainstorm** new approaches.

Source: CSET Analysis

Scientifically naive users have no or very basic scientific knowledge or hands-on laboratory experience. Regardless of access to AI, these individuals are more likely to produce a known, existing pathogen or toxin than a new or modified one. They can find information from publicly-available internet sources, or may use a chatbot to facilitate information-gathering. Importantly, the *perception* that a chatbot lowers barriers may be meaningful on its own if it engages new groups of actors who previously assumed that bioweapons were out of their reach.[6]

Scientifically knowledgeable users understand scientific concepts and research techniques. These individuals do not need AI to produce known, existing pathogens and toxins or modify them to be more dangerous. These users could use specialized AI tools, called biological design tools (BDTs), to design more severe, targeted, or dangerous biothreats, or to evade screening and detection measures.[7] Compared to traditional hypothesis-based bioengineering, BDTs may be more efficient and decrease laboratory work by prioritizing the most promising candidates for real-life testing.

Analysis

Our analysis considers two questions:

1. What biothreats exist without AI for both scientifically naive and scientifically knowledgeable users?
2. What specific capabilities change based on AI tools like AI chatbots and AI-based biological design tools (BDTs)?

Note that our analysis is based on the current state of biological AI tools in the Fall of 2023 and does not explore what future, more capable AI tools might be.

What biothreats exist without AI?

Malicious actors can already access the necessary information and resources to cause biological harm. This could involve acquiring or producing a known, existing pathogen or toxin, or developing a new one that has never been seen before. An individual's pre-existing level of scientific knowledge influences which outcome is most likely.

Scientifically Naive Users

Scientific novices can learn basic scientific concepts, techniques, and experimental protocols from online resources. These topics are widely taught and accessible because they are foundational scientific concepts. However, these same basic concepts could also be misused to cause harm.

Table 2 summarizes just some of the scientific topics that individuals can learn in theory or in practice from a variety of beginner-friendly sources. It is important to recognize that both information and laboratory experience are necessary to complete the long, multi-step process to produce a biological agent. Not all of these examples involve training in hands-on laboratory-based steps, which could limit a scientifically naive, malicious actor's success in producing a viable biorisk.

Table 2. Information Sources for Common Scientific Topic Areas

Scientific Topic Area	AP Biology	iGEM*	Biomanufacturing Workforce and Training**	Internet Resources
Bacterial Cell Biology				
Mammalian Cell Biology				
Viral Production		 ***		
Protein Expression and Purification				
DNA Cloning, Assembly, and Production				
Biosafety				

 indicates that individuals learn the underlying theory and principles

 indicates that individuals receive hands-on training performing a scientific procedure

* Specific projects may use some or all of the listed techniques and concepts.

** The specific training and competencies that a biomanufacturing operator receives depends on the pharmaceutical product type and manufacturing process.

*** Requires additional permission from the iGEM safety committee, and some viruses are off-limits.

Source: CSET Analysis

- High-school students who take **Advanced Placement Biology (AP Bio)** learn foundational biology concepts and techniques. Students learn foundational concepts, including the chemical and genetic basis of DNA, bacterial and viral genetic variation, and genetic engineering techniques.[8] Students also complete hands-on laboratory exercises that include genetically engineering *E. coli* with antibiotic resistance and aligning DNA sequences with BLAST software.[9]

- The **International Genetically Engineered Machine (iGEM) Competition** is a global synthetic biology competition for high school through post-grad teams.[10] Participants learn to design, clone, and assemble DNA "parts" to build and test a living biological system.[11] The competition aims to engage the next generation of synthetic biologists and capitalize on the future synthetic biology market, which experts predict will range from $37 billion to $100 billion by 2030.[12] Over 75,000 participants have competed in the competition.[13]

- **Biological and pharmaceutical manufacturing workers** learn chemistry and biology through certificate programs or associate degrees.[14] Typical job activities include operating laboratory equipment, preparing reagents, growing and maintaining living cells, and testing biological samples.[15] Not only do biomanufacturing workers learn these skills, but they gain the technical expertise to perform them to strict quality standards. This key workforce is facing shortages, and will need to be expanded to safeguard vulnerable U.S. medicine supply chains.[16]

- **Internet Resources:**

    - **AP Biology Help**: Hundreds of thousands of high-school students take AP Biology each year, creating a market for free, high-quality educational resources to help students pass the AP exam.[17] These resources are aimed at beginner-to-intermediate learners and are communicated in high-school-level language. Online courses and YouTube videos describe scientific concepts and walk viewers through the steps of a genetic engineering experiment.

    - **Laboratory Introductions and Online Courses:**[18] Some universities and laboratories develop introductory videos to teach students the basics before entering the lab for the first time. Others are developed by equipment and reagent manufacturers to help customers properly use their products. Topics range from teaching basic laboratory techniques

and equipment, like how to use a pipette or pH meter, to performing a DNA restriction digest or bacterial transformation.

In sum, the barrier to scientific information is not high, even for beginners. However, the feasibility of actually producing these agents depends on having both information and hands-on production success. Scientifically naive users are more likely to focus on known, existing pathogens for which information about how to produce or access them is readily available. Making new or modified pathogens, on the other hand, requires a deeper understanding of concepts like virology, immunology, and scientific experimentation.

Scientifically Knowledgeable Users

Experienced scientists can produce both existing and modified pathogens and toxins using scientific literature and previous laboratory experience, without AI tools. These individuals may already be proficient in basic laboratory techniques—like bacterial and mammalian cell culture, DNA cloning, and virus production—that could be misused.

These users can recreate known, existing pathogens in the lab using scientific literature, which provides technical information and step-by-step instructions to carry out scientific protocols. These publications are detailed and comprehensive to ensure reproducibility, and often include brands and catalog numbers for specific materials. This literature is prevalent and abundant, because many of the methods that could be envisioned in a misuse scenario are also fundamental, everyday laboratory techniques.*

Researchers can also manipulate pathogens and toxins to give them new functions, and regularly do so for a variety of scientific reasons. “Gain-of-function” research that enhances a pathogen’s ability to cause disease is widespread and has been conducted for decades using a variety of methods.[19]

* For example, reverse genetics is a way to obtain infectious virus particles using DNA or RNA that codes for viral genes. Protocols describing reverse genetics are well-described, given that reverse genetics was first used to rescue infectious poliovirus in 1981, yellow fever virus in 1989, and influenza A virus in 1999.

What specific capabilities change with AI?

Although AI tools are not required for biological misuse, they can alter the current and future risk landscape. The impact is not uniform and should be considered based on specific actors and scenarios of misuse. Here, we will outline some of these scenarios by assessing two types of AI tools—chatbots and biological design tools (BDTs)—in the context of other risk factors including the user's scientific expertise, the novelty of the biothreat, and user intent.

Chatbots

Chatbots generate text responses to users' questions by predicting plausible combinations of words based on natural language training data.[20] In a biological context, they might describe scientific concepts or provide links or references to additional resources, including those that could be misused. The impact on risk changes depending on the following three factors:

Scientific Expertise

Scientifically naive users may use chatbots to make the information-gathering process easier. Chatbots can explain scientific concepts, provide step-by-step instructions, and translate these materials into plain language. Without a chatbot, users must track down information, decide between multiple sources, and synthesize these pieces of information into a comprehensive plan. This process may be time-intensive, and there will be a learning curve as users learn basic biological concepts. In contrast, chatbots may facilitate the process by removing the need for users to know scientific terminology. This could enhance the allure of a chatbot as a "research assistant" and potentially increase the likelihood, ease, or speed with which a user identifies the necessary information to cause harm.

An initial assessment showed that chatbots can identify potential pandemic pathogens and methods to acquire them.[21] However, it is unclear whether or by how much chatbots provide an advantage over an internet search for existing resources. Notably, chatbot responses are not verified for factual accuracy, so it is also unclear whether chatbot-generated content is complete or accurate enough to result in real-world risk.

In contrast, scientifically knowledgeable users do not need a chatbot to translate technical jargon or introduce basic scientific concepts. Instead, experts may be more likely to use chatbots to speed up routine tasks, like identifying specialized scientific

literature or providing math or operational advice. While this could make some tasks easier or faster, it does not fundamentally change this user's core capabilities.

Biothreat Novelty

Chatbots are best-suited to provide resources for known, existing pathogens or toxins because they are trained on information that already exists. The wealth of highly-detailed scientific literature related to virology, immunology, epidemiology, and public health increases the likelihood that a chatbot could point users towards an existing protocol or generate a plausible set of instructions based on published procedures that have been done before.

Chatbots could also contribute to the development of novel pathogens or toxins by spurring conceptualization or idea generation. For example, a user intent on doing harm might request lists of immune targets or genes that contribute to pathogenicity and use this information to brainstorm potential risk-enhancing modifications.

User Intent

In addition to the intentional harms described above, chatbots can cause unintended risk if they suggest experimental plans without the proper safety procedures or that are scientifically unfounded. For example, chatbots may leave out the requirement for specific PPE or biocontainment infrastructure, combine steps from multiple sources that are not intended to be performed together, or hallucinate protocols altogether. Additionally, chatbots may generate text that is technically accurate but lacks appropriate scientific context.[22]

Biological Design Tools (BDTs)

Biological Design Tools (BDTs) engineer, predict, or simulate biological molecules, processes, or systems.[23] Researchers use them to predict physical characteristics, generate new types of biomolecules, and understand how biological factors—like genes, proteins, pathogens, and their hosts—interact. BDTs have already contributed to a wide range of biomedical, pharmaceutical, and basic research applications. Models can help researchers to design and optimize protein-based therapies, identify emerging viral strains that are likely to evade pre-existing immunity, design vaccine candidates that are less likely to be rendered obsolete when pathogens do evolve, and interpret DNA sequences and their impact on biological systems.[24] The first cohort of AI-generated drugs is currently in clinical trials, enabled by models that identified novel disease proteins and generated molecules to target them.[25]

While these applications are designed for beneficial purposes, they could be misapplied to cause harm. Instead of using BDTs to design therapies, malicious actors may use them to predict pathogens with more severe, targeted, or dangerous phenotypes, or that evade screening and detection measures. In some cases, the same models can be used for both beneficial and harmful purposes by recalibrating them to reward traits—like toxicity—that they are originally designed to minimize.[26] Different types of BDTs can contribute to harmful applications throughout the biothreat development process.*

Scientific Expertise

In their current state, BDT use requires both scientific and programming expertise. Users begin by deciding on a biological system to target, choosing the appropriate BDT, and setting the relevant biophysical parameters. Each of these steps requires a conceptual understanding of molecular pathways and how they can be impacted by genetic, structural, functional, or chemical modifications. Once the BDT has generated a set of predictions, users must be able to compare multiple options, physically produce the predicted biomolecule in a laboratory, and conduct the relevant experimental testing.

Scientific experts are most likely to be able to use a BDT successfully, and can benefit from these tools at each phase of the design-build-test-learn cycle.[27] Compared to traditional hypothesis-based bioengineering, BDTs can identify large-scale biological patterns and make predictions that would have been time-consuming or impossible for a researcher to identify by hand. High-quality predictions, combined with computational modeling for evaluation, can streamline the prioritization of lead candidates to accelerate further experimental testing.

At present, scientific novices are unlikely to use BDTs successfully. These tools would be more accessible to beginners if they did not require users to have a thorough understanding of the underlying scientific concepts or programming expertise. This could be achieved if chatbots could interpret plain-language requests, choose and program the appropriate BDT, translate technical results back into plain language, and provide a step-by-step laboratory protocol or control laboratory equipment to produce

* See Rose and Nelson 2023 for a more comprehensive analysis of subcategories of AI-enabled biological tools that contribute to biorisk and their relevant technological maturity levels: "Understanding AI-Facilitated Biological Weapon Development." The Centre for Long-Term Resilience, October 2023. https://www.longtermresilience.org/post/report-launch-examining-risks-at-the-intersection-of-ai-and-bio.

the result. While such chatbot "interpreters" have shown initial success for some tasks, they are not yet mature enough to fully enable a scientific beginner.*[28]

Biothreat Novelty

BDTs can contribute to the development of both known and novel biothreats. Malicious actors may use BDTs to facilitate the production of known, existing pathogens or toxins, or the design of novel or modified pathogens or toxins.

Although known, existing pathogens and toxins can already be produced with conventional methods, BDTs could increase risk if they are used to increase production efficiency or to redesign a biomolecule to evade screening measures. These could include models that are designed to increase biomanufacturing yield, for example by optimizing DNA sequences or growth conditions to maximize yield.[29] Alternatively, BDTs could potentially help malicious actors to avoid detection, for example by generating a protein sequence that is predicted to be functionally similar to a regulated toxin but has a different genetic code to bypass sequence-based screening.[30]

BDTs may have the greatest impact on risk if they are employed with the goal of designing novel or modified pathogens or toxins. For example, BDTs could enhance risk if they are used to identify immune targets, uncover genes that contribute to pathogenicity, or predict modifications that allow pathogens to infect new hosts or that enhance stability, transmissibility, or resistance to therapies. AI tools that generate protein designs, predict pathogen features or host-pathogen interactions, model the immune system, and simulate experiments could contribute to these outcomes and are in varying stages of development.[31]

User Intent

BDT predictions do not always function as intended when produced in the real world, and a pathogen or toxin may increase in risk when designed to do the opposite. Several underlying factors contribute to discrepancies between model predictions and biological outcomes.

Biological systems depend on complex, dynamic, and interconnected networks of molecular pathways that behave differently in different contexts. This means that

* In published studies, chatbots paired with BDTs have either only provided instructions or accomplished individual steps in larger experimental workflows and still required biological expertise to finish, required extensive programming expertise, or focused on chemical synthesis tasks, which are relatively straightforward compared to biological experimentation and BDT use.

experiments can yield vastly different results depending on whether they are conducted in a test tube, cells, tissues, or living organisms. This variability makes it difficult to predict how alterations to an organism will function in practice, even with substantial preliminary data.* AI models trained on data from one type of experimental system may also fail or have unintended consequences when their predictions are applied to other systems.

Incomplete, biased, or otherwise unrepresentative training data can also cause BDT predictions to fail. If data is collected from only a handful of representative cell lines, reference genomes, or populations, the resulting models may not be accurate for other systems. Cell-free laboratory experiments, like binding assays or X-ray crystallography, do not always represent the complexity of living systems, and can vary between labs due to minor variations in methods, equipment, or reagents. These nuanced limitations can be lost when dealing with large datasets, leading to the risk of overinterpreting subsequent model results.

However, unreliable predictions would only translate to real-world risk if the biological agent is actually produced and escapes containment. This scenario could be mitigated by incorporating biosecurity into the research design process.

Concluding Thoughts

The diversity of the biorisk landscape highlights the need to clearly identify which scenarios and actors are of concern. If this step is not considered, future policies may fail to address precisely those scenarios and actors of most concern. For example, biosecurity initiatives that use federal research funding as the policy lever only address one type of potentially risky actor. This approach can mitigate unintentional risk during the course of federally-funded biological research, but is unlikely to prevent deliberate bioweapon production from lone malicious actors (because such actors are unlikely to seek such funding to support bioweapon development). If lone malicious actors are of concern, then they will need to be targeted with different policy tools.

* In a previous CSET analysis of gain- and loss-of-function research, we encountered a study where a mutation strengthened a virus in cell culture but weakened it in living organisms. As another example, over 90% of new drugs fail clinical trials in patients even when results from laboratory and animal experiments look promising. See: Schuerger et al. "Understanding the Global Gain-of-Function Research Landscape." Policy Brief. Center for Security and Emerging Technology, August 2023. https://cset.georgetown.edu/publication/understanding-the-global-gain-of-function-research-landscape/; Sun, Duxin. "90% of Drugs Fail Clinical Trials," March 12, 2022. https://www.asbmb.org/asbmb-today/opinions/031222/90-of-drugs-fail-clinical-trials.

It will also be important to consider AI-enhanced risk within the current biorisk landscape. Both experts and non-experts can cause biological harm without the need for AI tools, highlighting the need for layered safeguards throughout the biorisk chain. Strategies that evaluate both AI-enhanced and AI-agnostic risks can differentiate between pre-existing risks and novel ones. This will be critical to build an effective foundation for biosecurity and biosafety oversight and more targeted measures to safeguard against AI-enabled risk.

As the United States revisits its biosecurity and biosafety oversight frameworks, a comprehensive review of the biorisk landscape could help to avoid ineffective policies that do not address the scenario of concern, or overbearing policies that hinder beneficial applications. By clearly defining the threats of concern and developing targeted mitigation measures, future policy can safeguard against the next generation of emerging biothreats.

Authors

Steph Batalis is a biotechnology research fellow at CSET.

Acknowledgments

The author would like to thank Igor Mikolic-Torreira and Matthew E. Walsh for their comprehensive review and Tessa Baker for her assistance and editorial support.

Endnotes

[1] AI's impact on biosecurity is emphasized in President Biden's Executive Order on AI, the UK's AI Safety Summit, and congressional testimonies. See: Exec. Order No. 14110, 88 FR 75191 (2023).; Department for Science, Innovation & Technology. "Capabilities and Risks from Frontier AI." Discussion Paper. AI Safety Summit 2023, October 24, 2023. https://assets.publishing.service.gov.uk/media/65395abae6c968000daa9b25/frontier-ai-capabilities-risks-report.pdf; De Vynck, Gerrit. "AI Leaders Warn Congress That AI Could Be Used to Create Bioweapons." *Washington Post*, July 25, 2023. https://www.washingtonpost.com/technology/2023/07/25/ai-bengio-anthropic-senate-hearing/.

[2] Exec. Order No. 14110, 88 FR 75191 (2023).; Batalis, Steph, and Vikram Venkatram. "Breaking Down the Biden AI EO: Screening DNA Synthesis and Biorisk." *Center for Security and Emerging Technology*, November 16, 2023. https://cset.georgetown.edu/article/breaking-down-the-biden-ai-eo-screening-dna-synthesis-and-biorisk/.

[3] National Institutes of Health. "Framework for Guiding Funding Decisions about Proposed Research Involving Enhanced Potential Pandemic Pathogens." 2017. https://www.phe.gov/s3/dualuse/Documents/P3CO.pdf.; The United States Government. "United States Government Policy for Oversight of Life Sciences Dual Use Research of Concern," March 29, 2012. https://www.phe.gov/s3/dualuse/Documents/us-policy-durc-032812.pdf.; The United States Government. "United States Government Policy for Institutional Oversight of Life Sciences Dual Use Research of Concern," September 25, 2015. https://www.phe.gov/s3/dualuse/Documents/durc-policy.pdf.

[4] Kelsey Lane Warmbrod, Michael G Montague, and Gigi Kwik Gronvall. "COVID-19 and the Gain of Function Debates." EMBO Reports 22, no. 10 (October 5, 2021): https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8490979/.; United States Government Accountability Office. "Public Health Preparedness: HHS Could Improve Oversight of Research Involving Enhanced Potential Pandemic Pathogens." Report to Congressional Committees. Washington, D.C.: United States Government Accountability Office, January 18, 2023. https://www.gao.gov/products/gao-23-105455.

[5] The National Science Advisory Board for Biosecurity (NSABB) offered recommendations to revise these policies in March 2023, and the Office of Science and Technology Policy (OSTP) issued a Request for Information to solicit public input regarding potential changes in September 2023.

[6] Headlines like the following tout the claim that chatbots will give untrained actors access to bioweapons within the next 3 years: "Friendly AI chatbots will be designing bioweapons for criminals 'within years'. Watch out, mad scientists – your careers are over."

[7] Sandbrink, Jonas B. "Artificial Intelligence and Biological Misuse: Differentiating Risks of Language Models and Biological Design Tools." arXiv, October 29, 2023. http://arxiv.org/abs/2306.13952.

[8] The coursework covers genetic engineering techniques including electrophoresis, PCR, recombinant DNA, gene cloning, bacterial transformation, and DNA sequencing (Unit 6.8, Biotechnology), the physiological and chemical basis of DNA mutations, and how bacteria and viruses accumulate mutations (Unit 6.7, Mutations). The level of detail is fairly in-depth; for example, the teacher manual states that

"when given a sequence of DNA containing a designated mutational change, students can predict the effect of the mutation on the encoded polypeptide and propose a possible resulting phenotype." See: College Board. "AP® Biology Course and Exam Description," 2020. https://apcentral.collegeboard.org/media/pdf/ap-biology-course-and-exam-description.pdf.

[9] College Board. "AP Biology Investigative Labs: An Inquiry-Based Approach." Teacher Manual. New York, NY, 2019. https://apcentral.collegeboard.org/media/pdf/ap-biology-teacher-lab-manual-effective-fall-2019_1.pdf.

[10] "iGEM Competition." Accessed November 14, 2023. https://competition.igem.org/.

[11] Top-scoring high school teams from recent years include:

- A group that designed two fusion proteins, assembled the corresponding DNA plasmids via molecular cloning and Gibson assembly, and expressed the fusion proteins and measured their function in mammalian cells.
- A group that expressed, produced, and purified a collagen-like protein in yeast and assembled it into thread, gel, and felt biomaterials. Notably, the team could not order their plasmids from commercial vendors because the sequence's repeating triplets were difficult to synthesize, so they designed and successfully employed a custom cloning and assembly strategy.

[12] Todd Kuiken. "Synthetic/Engineering Biology: Issues for Congress." Congressional Research Service, September 30, 2022. https://crsreports.congress.gov/product/pdf/R/R47265.; Warmbrod, Kelsey Lane, Marc Trotochaud, and Gigi Kwik Gronvall. "iGEM and the Biotechnology Workforce of the Future." *Health Security* 18, no. 4 (August 2020): 303–9. https://www.liebertpub.com/doi/10.1089/hs.2020.0017.

[13] "iGEM." Accessed December 8, 2023. https://igem.org/.

[14] Bio Jobs Hub. "Manufacturing Associate: Biotech Career Blueprint," January 6, 2022. https://biojobshub.wraltechwire.com/manufacturing-associate/.

[15] Catalyst, a Blog from WPI Graduate Studies. "Building Skills for a Successful Career in Biomanufacturing," January 31, 2023. https://wp.wpi.edu/catalyst/2023/01/31/building-skills-for-a-successful-career-in-biomanufacturing/.; Bio Jobs Hub. "Manufacturing Associate: Biotech Career Blueprint," January 6, 2022. https://biojobshub.wraltechwire.com/manufacturing-associate/.

[16] Batalis, Stephanie, and Anna Puglisi. "A Shot of Resilience: A Critical Analysis of Manufacturing Vulnerabilities in Vaccine Production." Policy Brief. Center for Security and Emerging Technology, May 2023. https://cset.georgetown.edu/publication/a-shot-of-resilience/; Office of Science and Technology Policy. "Building the Bioworkforce of the Future." Action Plan, June 2023. https://www.whitehouse.gov/wp-content/uploads/2023/06/Building-the-Bioworkforce-of-the-Future.pdf.

[17] In 2022, 237,338 students took the AP Bio Exam and 161,185 passed (score 3 or higher). See: College Board. "Student Score Distributions: AP Exams - May 2022," 2022. https://apcentral.collegeboard.org/media/pdf/ap-score-distributions-by-subject-2022.pdf.

[18] Examples of courses and YouTube videos that provide relevant biological training:

- An accurate and detailed animated YouTube video with over 2 million views describing the chemical basis of DNA mutations
- YouTube videos that show the genetic engineering laboratory module, step-by-step in real-time
- Khan Academy's completely free 37-unit, multi-module course
- Commercial bacterial genetic engineering kit
- 30-video YouTube playlist from Bio-Rad Laboratories, a major equipment supplier
- YouTube videos from top suppliers showing how to perform cell transfection and site-directed mutagenesis
- Learn to pipette in a simulated lab.

[19] Schuerger et al. "Understanding the Global Gain-of-Function Research Landscape." Policy Brief. Center for Security and Emerging Technology, August 2023. https://cset.georgetown.edu/publication/understanding-the-global-gain-of-function-research-landscape/.

[20] Dunham, James. "Large Language Models (LLMs): An Explainer." *Center for Security and Emerging Technology*, July 31, 2023. https://cset.georgetown.edu/article/large-language-models-llms-an-explainer/.

[21] Soice, Emily H., Rafael Rocha, Kimberlee Cordova, Michael Specter, and Kevin M. Esvelt. "Can Large Language Models Democratize Access to Dual-Use Biotechnology?" arXiv, June 6, 2023. https://arxiv.org/abs/2306.03809.

[22] A recent study used an LLM to suggest a novel anticancer therapy. The model recommended a plant-based cannabinoid scaffold; the authors note that this selection might be "an artifact of the abundance of unverified information online on cures based on cannabinoids." See Boiko, Daniil A., Robert MacKnight, and Gabe Gomes. "Emergent Autonomous Scientific Research Capabilities of Large Language Models." arXiv, April 11, 2023. http://arxiv.org/abs/2304.05332.

[23] Sandbrink, Jonas B. "Artificial Intelligence and Biological Misuse: Differentiating Risks of Language Models and Biological Design Tools." arXiv, October 29, 2023. http://arxiv.org/abs/2306.13952.

[24] Many studies have focused on the mentioned applications, and their contributions are not all possible to mention here. A select, non-comprehensive list is included in this citation. Shanehsazzadeh, Amir, Sharrol Bachas, George Kasun, John M. Sutton, Andrea K. Steiger, Richard Shuai, Christa Kohnert, et al. "Unlocking de Novo Antibody Design with Generative Artificial Intelligence." bioRxiv, January 9, 2023. https://www.biorxiv.org/content/10.1101/2023.01.08.523187v1; Hie, Brian L., Varun R. Shanker, Duo Xu, Theodora U. J. Bruun, Payton A. Weidenbacher, Shaogeng Tang, Wesley Wu, John E. Pak, and Peter S. Kim. "Efficient Evolution of Human Antibodies from General Protein Language Models." *Nature Biotechnology*, April 24, 2023, 1–9. https://www.nature.com/articles/s41587-023-01763-2; Pandi, Amir, David Adam, Amir Zare, Van Tuan Trinh, Stefan L. Schaefer, Marie Burt, Björn Klabunde, et al. "Cell-Free Biosynthesis Combined with Deep Learning Accelerates de Novo-Development of Antimicrobial Peptides." *Nature Communications* 14, no. 1 (November 8, 2023): 7197. https://www.nature.com/articles/s41467-023-42434-9.; Thadani, Nicole N., Sarah Gurev, Pascal Notin, Noor Youssef, Nathan J. Rollins, Daniel Ritter, Chris Sander, Yarin Gal, and Debora S. Marks. "Learning from Prepandemic Data to Forecast Viral Escape." *Nature* 622, no. 7984 (October 2023): 818–25.

https://www.nature.com/articles/s41586-023-06617-0.; Zvyagin, Maxim, Alexander Brace, Kyle Hippe, Yuntian Deng, Bin Zhang, Cindy Orozco Bohorquez, Austin Clyde, et al. "GenSLMs: Genome-Scale Language Models Reveal SARS-CoV-2 Evolutionary Dynamics," November 23, 2022. doi:10.1101/2022.10.10.511571.; Gema, Aryo Pradipta, Michał Kobiela, Achille Fraisse, Ajitha Rajan, Diego A. Oyarzún, and Javier Antonio Alfaro. "Vaxformer: Antigenicity-Controlled Transformer for Vaccine Design Against SARS-CoV-2." arXiv, May 18, 2023. http://arxiv.org/abs/2305.11194.; Ji, Yanrong, Zhihan Zhou, Han Liu, and Ramana V Davuluri. "DNABERT: Pre-Trained Bidirectional Encoder Representations from Transformers Model for DNA-Language in Genome." *Bioinformatics* 37, no. 15 (August 9, 2021): 2112–20. https://academic.oup.com/bioinformatics/article/37/15/2112/6128680.; Vaishnav, Eeshit Dhaval, Carl G. de Boer, Jennifer Molinet, Moran Yassour, Lin Fan, Xian Adiconis, Dawn A. Thompson, Joshua Z. Levin, Francisco A. Cubillos, and Aviv Regev. "The Evolution, Evolvability and Engineering of Gene Regulatory DNA." *Nature* 603, no. 7901 (March 2022): 455–63. https://www.nature.com/articles/s41586-022-04506-6.; https://www.embopress.org/doi/full/10.15252/msb.20199083.

[25] Arnold, Carrie. "Inside the Nascent Industry of AI-Designed Drugs." *Nature Medicine* 29, no. 6 (June 1, 2023): 1292–95. https://www.nature.com/articles/s41591-023-02361-0.

[26] Kuiken, Todd. "Artificial Intelligence in the Biological Sciences: Uses, Safety, Security, and Oversight." Congressional Research Service, November 22, 2023. https://crsreports.congress.gov/product/pdf/R/R47849.

[27] Rose, Sophie, and Cassidy Nelson. "Understanding AI-Facilitated Biological Weapon Development." The Centre for Long-Term Resilience, October 2023. https://www.longtermresilience.org/post/report-launch-examining-risks-at-the-intersection-of-ai-and-bio.

[28] Boiko, Daniil A., Robert MacKnight, and Gabe Gomes. "Emergent Autonomous Scientific Research Capabilities of Large Language Models." arXiv, April 11, 2023. http://arxiv.org/abs/2304.05332.; Inagaki, Takashi, Akari Kato, Koichi Takahashi, Haruka Ozaki, and Genki N. Kanda. "LLMs Can Generate Robotic Scripts from Goal-Oriented Instructions in Biological Laboratory Automation." arXiv, April 18, 2023. https://arxiv.org/abs/2304.10267.; Bran, Andres M., Sam Cox, Oliver Schilter, Carlo Baldassari, Andrew D. White, and Philippe Schwaller. "ChemCrow: Augmenting Large-Language Models with Chemistry Tools." arXiv, October 2, 2023. https://arxiv.org/abs/2304.05376.

[29] Rose, Sophie, and Cassidy Nelson. "Understanding AI-Facilitated Biological Weapon Development." The Centre for Long-Term Resilience, October 2023. https://www.longtermresilience.org/post/report-launch-examining-risks-at-the-intersection-of-ai-and-bio.

[30] Walsh, Matthew, Stephanie Batalis, Caroline Schuerger, and Gigi Gronvall. "Safeguarding Mail-Order DNA Synthesis in the Age of Artificial Intelligence." *Applied Biosafety*. (Forthcoming, in publication).

[31] See Rose and Nelson 2023 for subcategories of AI-enabled tools that contribute to these applications and their relative maturity.