AUGUST 2025

Opportunities to Strengthen U.S. Biosecurity from AI-Enabled Bioterrorism

What Policymakers Should Know

AUTHORS
Georgia Adamson
Gregory C. Allen

A Report of the CSIS Wadhwani AI Center

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES | WADHWANI AI CENTER

Table of Contents

AUGUST 2025

Opportunities to Strengthen U.S. Biosecurity from AI-Enabled Bioterrorism

What Policymakers Should Know

By Georgia Adamson and Gregory C. Allen

> *"AI will unlock nearly limitless potential in biology. . . . At the same time, it could create new pathways for malicious actors to synthesize harmful pathogens and other biomolecules."*
>
> ***— America's AI Action Plan, Trump Administration, 2025***

> *"To protect against adversaries' misuse of biotechnology, the United States must be able to detect and characterize the widest possible array of biological threats and do so early."*
>
> ***—2025 National Security Commission on Emerging Biotechnology Report***

Executive Summary

The falling barriers to bioterrorism are set to accelerate in the emerging age of AI and biotechnology.

The convergence of artificial intelligence (AI) and biotechnology presents incredible opportunities to accelerate scientific discovery and improve global health—opportunities that the United States must seize. At the same time, the dual-use nature of AI-enabled biotechnology poses new security challenges for policymakers worldwide.

The cost and technical expertise required to develop bioweapons have been sharply declining for nearly a century. Now, rapid advancements in AI capabilities present two emerging biosecurity threats in which AI could assist actors in developing bioweapons:

1. **Popular large language models (LLMs) could soon drastically lower the informational barriers for planning and executing biological attacks.** Recent assessments of LLMs and

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES | WADHWANI AI CENTER

other commercial AI capabilities indicate that models are "**on the cusp**" of meaningfully helping novices develop and acquire bioweapons by providing critical information and step-by-step guidance.

2. **Future AI biological design tools (BDTs) could assist actors in developing more harmful or even novel epidemic- or pandemic-scale pathogens.** **Rapid advancements** in state-of-the-art BDTs—illustrated by the foundation model **Evo 2**—point to a world in which **more capable** models could help develop new or enhanced pathogens and evade existing safeguards.

Current U.S. biosecurity measures are ill-equipped to prevent AI-enabled biological threats.

Because no single intervention is sufficient to address the complexity, diversity, and evolving nature of biological threats, successfully mitigating biological risks from AI will require multiple **layers of defense** spread across multiple steps of the bioweaponization pathway. Implementing AI-model safeguards and strengthening oversight of nucleic acid synthesis represent two critical choke points in AI-enabled bioweapons development and thus opportunities for targeted policymaking. However, these measures face tough challenges ahead:

1. **Critical safeguards in advanced BDTs can already be circumvented by users post-deployment.** The first Evo model **shows** that users are already able to bypass key BDT security measures by adding harmful viral data originally removed from the model's training dataset back into it later via fine-tuning. Concerningly, the Evo series represents one of only a **handful** of BDTs that possess any such safeguards at all.

2. **List-based approaches to screening nucleic acid synthesis will likely fail to detect future AI-generated sequences that do not match known agents and toxins.** To date, there are only 63 regulated toxins, pathogens, and bacteria on the **Select Agents and Toxins List** compiled by the Department of Health and Human Services and Department of Agriculture. When including export-controlled items on the Bureau of Industry and Security's **Commerce Control List** (CCL) for international orders, there are only approximately 150 unique items. Special safeguards apply when a customer requests synthesis of any of these regulated agents and toxins.

   In theory, synthesis providers' screening mechanisms work until a BDT creates the 64th (or, for international orders, the 151st) highly contagious and lethal organism that is not yet on one of these lists. There are thus serious risks that these lists of regulated sequences are an increasingly **inadequate** line of defense to catch new AI-generated ones. Static lists are insufficient for a world in which novel organisms are relatively easy to design. Though such capabilities are considerably beyond what the most advanced AI systems are capable of today, they may be closer than they seem given the demonstrated pace of technological acceleration. Ongoing efforts to expand screening protocols to capture additional sequences of concern are nascent and could quickly fall behind due to slow and uneven data collection methods.

The United States must strengthen biosecurity in the AI era.

As the Trump administration's July 2025 **AI Action Plan** rightly identifies, investing in biosecurity is an essential component to the United States' national security strategy. To pursue a high-return, low-cost biodefense policy, U.S. policymakers and lawmakers should undertake the following recommendations:

1. **Fund the National Institute of Standards and Technology (NIST) and the U.S. Center for AI Standards and Innovation (CAISI) to continue their critical work at the intersection of AI and biosecurity.** NIST and CAISI are conducting unique and urgently needed work to prevent the misuse of AI and biotechnology for bioterrorism. While both organizations are central to the success of the Trump AI Action Plan, the administration's proposed **cuts to their budgets** threaten these critical centers of government expertise, capacity, and industry engagement. U.S. policymakers and lawmakers should support NIST and CAISI with funding proportionate to scope of their responsibilities.
2. **Evaluate frontier BDT capabilities with testing led by CAISI and with support from the interagency Testing Risks for AI and National Security (TRAINS) Taskforce and the international network of AI Safety Institutes.** A comprehensive U.S. biosecurity strategy should include evaluating the unique capabilities and associated risks posed by BDTs, including non-commercial BDTs, as they evolve. CAISI should conduct BDT model evaluations to assess frontier BDTs for emerging high-impact biological capabilities, leveraging existing resources including the TRAINS Taskforce and international AI Safety Institute network to do so.
3. **Develop a standardized, AI-enabled, nucleic acid synthesis-screening system.** Current list-based synthesis-screening measures are uneven across industry and will likely fail to detect AI-generated agents over time. The White House Office of Science and Technology Policy (OSTP) and relevant U.S. agencies should develop a plan to integrate AI capabilities into standardized screening measures that can detect new and augmented AI-generated sequences beyond lists of federally regulated agents and toxins. This system could implement a tiered risk assessment of sequence orders based on sequence function and strengthen collective biosecurity while reducing compliance burdens for industry.

Introduction

Hollywood has depicted mad scientists causing global catastrophes with devious technologies for so long, the idea has its own **Wikipedia** entry. Unfortunately, in the emerging age of AI-enabled biotechnology, evil geniuses using technology for mass destruction could soon move from fiction to fact. Worse, such capabilities could also fall within reach of "evil morons" aided by genius-level AI.

Historically, one of the greatest gifts to U.S. national security has been that building weapons of mass destruction (WMDs) is expensive and complicated. There are, of course, plenty of simple, cost-effective ways to wreak havoc on localized groups of people. But thankfully, the enormous expertise and budget needed to develop and deploy WMDs such as nuclear bombs is prohibitively high for the average terrorist. Even well-resourced nation-states like Saddam Hussein's Iraq, which spent around **$10 billion** on its nuclear weapons program, **failed**.

It is in the United States' interest that WMDs continue to require lots of money and highly trained scientists. Unfortunately, the barriers to developing some WMDs, especially biological weapons, have fallen sharply over the last century. As Massachusetts Institute of Technology (MIT) professor Kevin Esvelt **told Congress** in 2022, "We live in the biological version of a world in which weapons-grade plutonium can be mail-ordered and thousands of engineers have the skills to assemble a nuclear device, but no one knows exactly which design would work." However, with rapid advancements in large language models and **biological design tools**, this final barrier may be increasingly **shrinking**.

> *Historically, one of the greatest gifts to U.S. national security has been that building weapons of mass destruction (WMDs) is expensive and complicated.*

U.S. policymakers face the challenge of strengthening biosecurity measures for an AI era while ensuring that biotechnology innovation can still flourish. AI-enabled **biotechnology** stands to improve billions of lives globally by accelerating vaccine discovery, improving healthcare diagnostics, and enhancing countless other fields. At the same time, the trajectory of AI's biological capabilities points to a world in which evading existing safeguards could become increasingly straightforward. As AI continues to revolutionize the biotechnology field, policymakers have the tough job of promoting its incredible benefits for science and society while preventing bad actors from exploiting these tools to engineer epidemic- or pandemic-capable agents.

The Trump administration has recently expressed commitment to investing in U.S. biosecurity in its July 2025 **AI Action Plan**, a comprehensive policy playbook outlining the administration's priorities for AI. The plan rightly identifies AI's dual-use capabilities in the biological domain and the need to mature the United States' biosecurity strategy to counter emerging threats from bad actors. Its recommendations for strengthening nucleic acid synthesis screening and developing government AI evaluation capabilities, while largely high-level at this stage, are nevertheless a promising start for updating the United States' biodefense for an AI era.

This paper surveys emerging AI-enabled bioterrorism risks and offers further recommendations for enhancing synthesis screening and advancing government AI evaluation capabilities for biothreats.[1] It examines two AI-enabled bioterrorism risks: (1) falling informational barriers to bioterrorism from LLMs and (2) harmful biological design risks from BDTs.[2] It then turns to two biosecurity measures: (1) BDT security controls and (2) list-based nucleic acid synthesis screening. While these measures are essential components of a multilayered U.S. biodefense system, the report outlines how rapid technological advancements are exposing critical gaps for meeting AI-enabled biorisks. The report concludes with three recommendations for U.S. policymakers and lawmakers.

The Falling Barriers to Bioweapons Development

The cost and technical expertise required to **develop bioweapons** have been sharply declining for nearly a century. During World War II, the U.S. bioweapons program (later **terminated** in 1969 during the Nixon administration) employed around **4,500** people, of whom a **significant** share were technical experts. Its multiyear budget totaled **$60 million**—close to 3 percent of total spending on the Manhattan Project, which was itself at its peak almost **0.5 percent** of U.S. GDP. Similarly, Japan's World War II biological warfare program employed roughly **3,000** people, including hundreds of Japan's top medical experts (and more than a few **psychopaths**).

Decades later, in the 1990s, the Japanese terrorist organization **Aum Shinrikyo** made multiple attempts to develop and deploy biological weapons using botulinum and anthrax. Thankfully, Aum Shinrikyo's efforts **failed**. However, the group **successfully executed** many of the steps for developing and delivering bioweapons despite having fewer than two dozen technical members and a budget substantially smaller than a nation-state (though still in the **millions** of dollars).[3] Aum Shinrikyo still needed smart people with advanced degrees, but far fewer than the United States or Japan required during World War II.

Then, in 2017, two **Canadian scientists** successfully recreated the previously extinct horsepox virus for just **$100,000**. Rather than developing the virus themselves, the scientists outsourced much of the initial work—custom ordering fragments of the DNA from a commercial synthesis lab, which "printed" and **shipped** the viral DNA back to them via mail.[4] The scientists then **linked** the fragments together in a lab and introduced them into cells using a helper virus, producing the final horsepox virus. Though the study exclusively **aimed** to improve vaccine and cancer treatments (rather than produce bioweapons), it nevertheless attracted **widespread alarm** at the time by indicating that reviving smallpox—a close cousin to horsepox and one of the deadliest diseases known to mankind—would, as ***Science*** reported, "probably take a small scientific team with little specialized knowledge half a year."

Preventing AI-Enabled Bioweapons as a Matter of Urgency

Today, AI threatens to reduce the number of experts needed to develop bioweapons further still. On March 5, 2025, a paper released by former Google CEO Eric Schmidt, former Scale AI CEO Alexandr Wang, and Center for AI Safety Executive Director Dan Hendrycks **argued** that

> AI could provide step-by-step guidance on designing lethal pathogens, sourcing materials, and optimizing methods of dispersal. What once required specialized knowledge and resources could become accessible to individuals with malevolent intent, dramatically increasing the potential for catastrophic outcomes.

Dario Amodei, CEO of the AI lab Anthropic, agrees and predicted in 2023 that such capabilities could be possible in as few as two to three years. Anthropic reissued a similar timeline earlier this year, and reported testing showing that one of its recent models, **Claude 3.7 Sonnet**, already "demonstrates concerning improvements in its capacity to support aspects of biological weapons development."

> *Even traditionally staunch opponents of AI regulations . . . have acknowledged that the AI-related risk from bioweapons deserves special attention and mitigation.*

Some, including U.S. Vice President **J.D. Vance**, have said that the United States should greet Big Tech's calls for AI regulation with skepticism, arguing that onerous government regulations squash startups while strengthening incumbent tech giants that can bear compliance costs. However, in the case of AI-enabled bioweapons, calls for greater security are grounded in good sense. Even traditionally **staunch opponents** of AI regulations such as Meta CEO Mark Zuckerberg have **acknowledged** that the AI-related risk from bioweapons deserves special attention and mitigation.

The first Trump administration, to its credit, also recognized that growing biosecurity risks were real and demanded action in its 2018 **National Biodefense Strategy,** which rightly stated, "Biological threats—whether naturally occurring, accidental, or deliberate in origin—are among the most serious threats facing the United States. . . . As we reap the benefits from biotechnologies, we must also understand and consider the risks they may pose." The strategy further pledged to "promote appropriate measures to impede misuse of life sciences and biotechnology" as part of the administration's strategic vision in which "the United States actively and effectively prevents . . . risk from natural, accidental, or deliberate biological threats."

The second Trump administration has recently reaffirmed this commitment in its highly anticipated **AI Action Plan**. The administration calls for the federal government to develop "new tools and infrastructure" for a "multi-tiered approach" to nucleic acid synthesis screening, including more stringent screening requirements and a data sharing mechanism for industry (discussed in later sections of this paper). It also encourages relevant agencies, including CAISI, to evaluate models for biological and other national security-related risks (see recommendation two of this paper).

Emerging Biosecurity Risks from AI

The Trump administration is right to identify biosecurity as a national security imperative in the AI era. However, the time to prepare for AI-enabled biological threats is quickly running out. Rapid advancements in AI capabilities present two **scenarios** in which AI could meaningfully assist actors with developing bioweapons. First, AI labs warn that popular commercial LLMs could soon drastically lower informational barriers to planning and executing biological attacks. Second, future BDTs could assist actors in producing more harmful or novel agents and toxins of epidemic or pandemic scale.

LARGE LANGUAGE MODELS: REDUCING INFORMATIONAL BARRIERS TO BIOTERRORISM

In February 2025, AI lab OpenAI published a **safety assessment** of **Deep Research**, one of its advanced AI capabilities, in which it stated, "Our evaluations found that deep research can help experts with the operational planning of reproducing a known biological threat . . . [and] indicate our models are on the cusp of being able to meaningfully help novices create known biological threats." The company underlined this conclusion in an April 2025 **safety assessment** of its o3 and o4-mini models, adding:

> We expect current trends of rapidly increasing [biological] capability to continue, and for models to cross this [high-risk] threshold in the near future. In preparation, we are intensifying our investments in safeguards. . . . At the same time, **we also encourage broader efforts to prepare for a world where the informational barriers to creating such threats are drastically lower.**

Less than a week after OpenAI published this assessment, a study—conducted by researchers from SecureBio, MIT, the Center for AI Safety, and Federal University of ABC in Brazil—revealed o3 outperformed **94 percent** of expert virologists on a subject matter test, which measured the ability to troubleshoot **complex** virology lab protocols. The results demonstrate that widely available LLMs such as o3 can provide expert-level advice on dual-use virology topics, prompting what the study called an "urgent need for thoughtful access controls" on these models.

OpenAI has recently implemented stricter biological safeguards in its new ChatGPT agent, a model capable of semi-autonomous, multi-step research and task execution. In a July 2025 assessment of the model, the lab **stated** that it would treat the agent as "highly capable" in the biological portion of its **Preparedness Framework**, a classification reserved for capabilities that significantly increase existing risks for severe harm. Though OpenAI **clarified** that the move was precautionary, it nevertheless reiterated that increasingly available AI bio capabilities could soon meaningfully help actors develop biological threats with fewer barriers to entry.

OpenAI is not the only lab reporting rapid advancements in its models' abilities to provide users with information critical to bioweapons development. Anthropic's assessment of **Claude 3.7 Sonnet** similarly revealed that the model "provides better advice in key steps of the weaponization pathway, makes fewer mistakes in critical steps . . . and ultimately make[s] solving complex problems [in WMD creation] faster" than previous models.

Of special note, in a trial that tested whether the model could meaningfully assist users in bioweapons planning and acquisition, one Anthropic employee achieved a score of 91 percent using Claude 3.7 Sonnet. The trial set 80 percent as the threshold in which a model is deemed to substantially increase the risk of **catastrophic misuse** (though Anthropic reported that the average test scores from all participants fell below this threshold). Anthropic reported that external red teaming of its more recent model, **Claude Opus 4**, demonstrates "more accurate and comprehensive answers in some areas of bioweapons-related topics" compared to Claude 3.7, leading the lab to **activate** the next level of its security control framework.

Safety evaluations of the R1 model developed by Chinese AI lab DeepSeek also demonstrate significant bio risks. **Testing** revealed that R1 can be tricked into producing content about harmful topics—a process known as "jailbreaking"—more easily than its Western AI counterparts due to fewer guardrails embedded within the model. In a **February 2025 interview**, Anthropic's Dario Amodei said his company's evaluation of DeepSeek found it generated information critical to producing bioweapons that "can't be found on Google or can't be easily found in textbooks," adding that "the DeepSeek model did the worst of basically any model we'd ever tested in that it had absolutely no blocks whatsoever against generating this information."

Such assessments from leading AI labs demonstrate that LLMs are rapidly approaching or even exceeding critical security thresholds for providing users key bioweapons development information—with some models already demonstrating capabilities that surpass expert-level knowledge. While some of the leading companies are voluntarily imposing safeguards, the overall trajectory nevertheless points toward a near-term future in which policymakers must confront bioterrorism risks not just from sophisticated state and terrorist organizations, but potentially from individuals with little technical background but access to popular LLMs.

BIOLOGICAL DESIGN TOOLS: GENERATING PATHOGENIC AND TOXIC SEQUENCES

On February 19, 2025, researchers from the nonprofit Arc Institute, Stanford University, and AI chip designer Nvidia **announced** the world's **largest** biological AI foundation model, known as Evo 2. The model (also known as a **biological design tool**, or BDT) is trained on a dataset containing over **128,000** genomes from an astonishingly **diverse set** of lifeforms, including humans, animals, plants, bacteria, and other organisms, in addition to certain viruses. This enormous database reportedly allows it to identify novel patterns between different DNA sequences and to design the genomes of entirely new organisms. Using this BDT, researchers can enter the sequences of real and/or hypothetical organisms to simulate how that organism would develop in the real world.

Evo 2 is already demonstrating impressive results in the biomedical field. The tool reportedly achieves **90 percent** accuracy in predicting whether various mutations of a breast cancer-related gene will be benign or possibly pathogenic. The implications of this high-fidelity simulation capability are enormous. According to the **Arc Institute**, "Insights like this could save countless hours and research dollars needed to run cell or animal experiments." Similarly, **Stanford University** writes, "Imagine being able to speed up evolution—hypothetically—to learn which genes might have a harmful or beneficial effect on human health. Imagine, further, being able to rapidly generate new genetic sequences that could help cure disease."

Evo 2 is just one of many BDTs making valuable contributions to science. Google DeepMind's **AlphaFold** model, for instance, won its developers the 2024 **Nobel Prize** in Chemistry for accurately predicting the three-dimensional structure of proteins from their amino acid sequences, greatly outperforming previous methods. Likewise, IBM's **Biomedical Foundation Models** are helping accelerate and streamline drug discovery by widening the scope for novel molecules and determining unsuitable molecules earlier in the research process. These tools and many others are already transforming numerous scientific fields and are extremely **promising** for the future of global health. They should first and foremost be understood in these terms.

However, if BDTs such as Evo 2 can generate good outputs, then presumably more capable future tools could create bad ones, too. Imagine, for instance, bird flu: Some strains have a reported human mortality rate of more than **50 percent**, meaning that approximately one out of every two people that contract the disease will die.[5] There are other versions of flu that are not as lethal but are more contagious. Someone who contracts the more common seasonal flu, for instance, will spread it to **1-2 other people** on average. Malicious actors, however, might use a more advanced BDT with generative capabilities to create a new version of bird flu that is both highly lethal and highly contagious. Such a scenario would likely make the Covid-19 pandemic—which has claimed more than **27 million** lives and shaken the world economy—look like the common cold.

For now, BDTs are incapable of designing such pathogens from scratch. In a December 2024 **report**, experts representing several leading universities, AI companies, and research organizations argued that the reality of AI-enabled biorisk is that these tools cannot currently be used for novel pathogen design due to a lack of viral training data as a principle bottleneck. A March 2025 **report** from the National Academies of Sciences, Engineering, and Medicine (NASEM) agreed, arguing that while today's state-of-the-art BDTs can design much simpler biological structures such as molecules, they are currently unable to design self-replicating pathogens, which are orders of magnitude more complex. As NASEM stated, it is "unlikely that currently available viral sequence data are sufficient to train such a model," and current BDTs still lack the requisite understanding of how complex biological systems interact to unlock such capabilities.

Drawing on publicly available literature, the December 2024 report **concluded** that there are "no known examples of current AI biological tools being misused to cause real-world harm" and that "the available literature does not support the notion that access to biological information and planning via current, publicly available LLMs can significantly increase biorisk." However, the authors admitted that "this does not offer conclusions for future models," and that further research is needed in this area.

But no serious approach to national security can accept an exclusive focus on risks identical to those that have already occurred. As the attacks of September 11, 2001, and Pearl Harbor painfully demonstrated, malicious actors—whether terrorists or nation-states—are innovative, opportunistic, and able to attack in unprecedented ways.

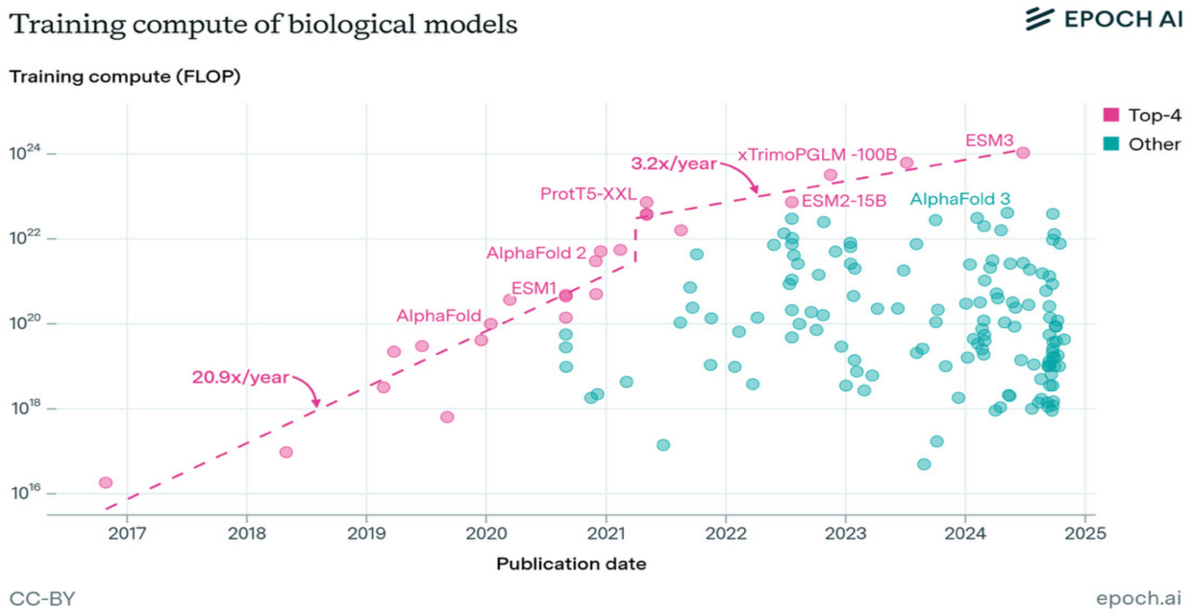
> *No serious approach to national security can accept an exclusive focus on risks identical to those that have already occurred.*

Indeed, as an illustration of just how fast this landscape is moving, the 2024 report cited a January 2024 **OpenAI** red-teaming exercise and a June 2024 safety assessment of Anthropic's **Claude 3** model as evidence that "current LLMs . . . do not pose an immediate [bio]risk." At the time, OpenAI and Anthropic reported little to no uplift in users' ability to plan and execute a biological attack compared to using other information sources such as the internet. However, both AI labs have since updated their risk assessments to flag substantial and immediate biorisks, with OpenAI **concluding** in April 2025 that it expects its models to demonstrate "**high risks**"—meaning that they could substantially increase the likelihood and frequency of bioterrorist attacks by providing meaningful assistance to novices—in the near future.

The Evo model series also demonstrates significant improvements in BDT capabilities over the course of just one year. Compared to its Evo 1 predecessor (released February 2024), Evo 2 contains **30 times** more data and can process more than eight times as much genetic information at a time. As a result, the model demonstrates substantially better generative, predictive, and biological design capabilities, representing in many instances **state-of-the-art** performance (see Appendix A for a comparison of Evo 1 and Evo 2 model specifications and capabilities).

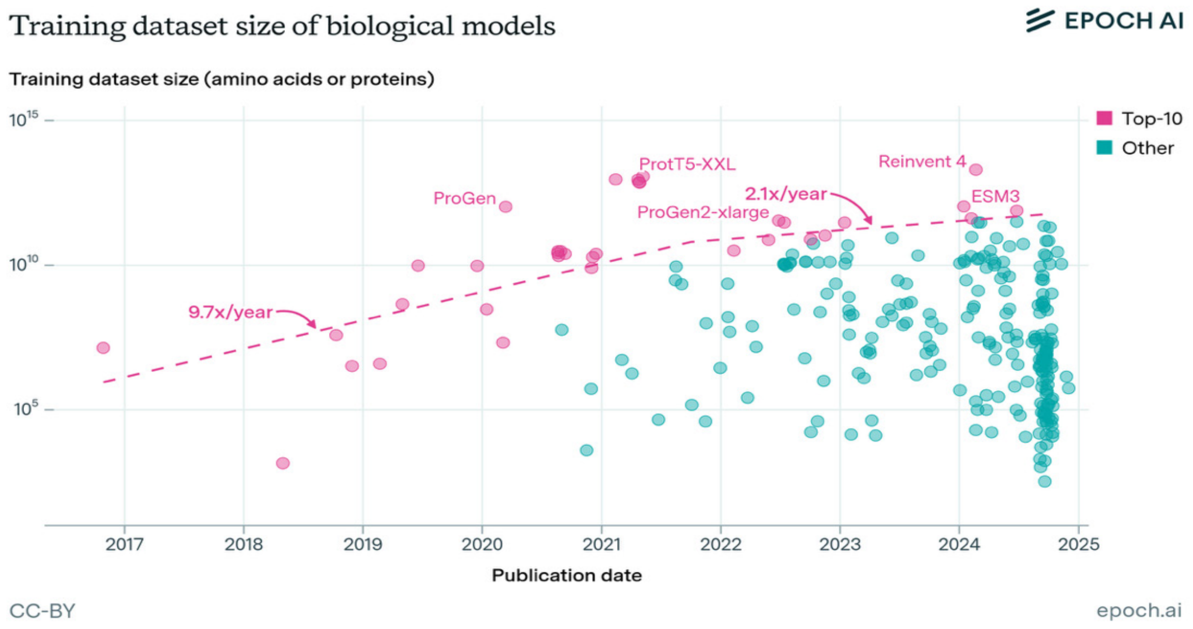
These rapid improvements can be observed across BDTs at scale. Epoch AI, a nonprofit research institute, estimates that the training compute of the top BDTs (a common metric for determining AI model performance) increased by an average of nearly **21 times per year** from 2017 to 2021, the equivalent to doubling just under every three months (Figure 1). The size of model training data, another metric for predicting future performance, grew nearly 10 times per year on average from 2017 to 2021 (Figure 2). Though growth in both of these metrics has since slowed to a more modest 2-3 times per year, BDTs are nevertheless demonstrating continued advancement with more impressive capabilities and thus dual-use concerns.

Figure 1: Training Compute of BDTs



Source: Epoch AI.

Figure 2: Size of BDT Training Datasets



Source: Epoch AI.

These rapid advancements in BDTs and LLMs suggest that current limitations do not preclude future AI models from unlocking further capabilities, including potentially much more harmful ones than current models possess. U.S. policymakers should not wait to examine the security implications of more capable models only after they have arrived.

As one biosecurity expert told CSIS, many studies rightly state that BDTs cannot yet design pandemic-scale pathogens or toxins, yet their emphasis on current limitations paints an overly conservative picture of the trajectory of future capabilities.[6] For example, this source argued, while data availability remains a key developmental bottleneck, experts are working to improve datasets that will inevitably solve this issue. **Companies** are already working to create biological datasets specifically for training more capable BDTs able to design novel organisms. (Additionally, the Trump administration's AI Action Plan has **tasked** federal agencies with exploring the creation of a whole-genome sequencing program that could generate enormous biological datasets to be used in training future models.) Given both public and private sector efforts to overcome bottlenecks to more capable models, BDT capabilities are likely to substantially improve in the next three-to-five years, the expert predicted.

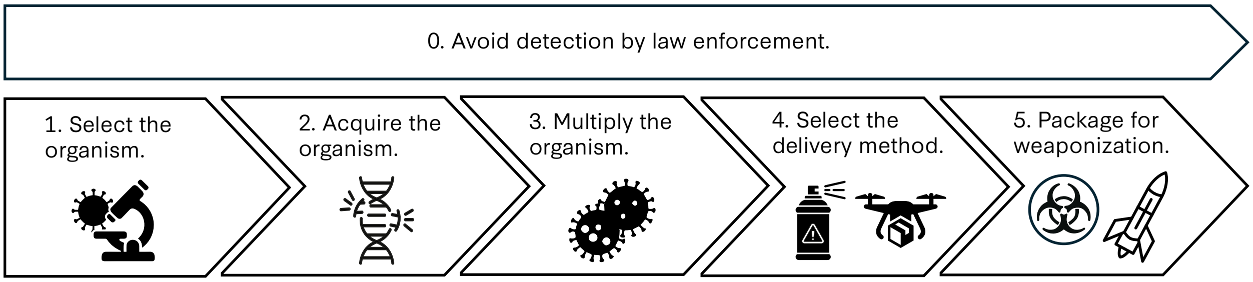
> *In an August 2024 paper, researchers from Johns Hopkins, Stanford, and Fordham Universities argued . . . "The essential ingredients to create highly concerning advanced biological models . . . may already exist or soon will. Establishment of effective governance systems now is warranted."*

Other scholars—including several members of the **team** behind Evo 2—agree that dual-use BDT capabilities are likely to accelerate in the near term. In an **August 2024 paper**, researchers from Johns Hopkins, Stanford, and Fordham Universities argued that while models are currently limited in their generation capabilities, "The rapid progress of AI—and the ever-larger resources being invested in computation and data generation—suggests that capabilities are likely to accelerate," and that "limitations are not likely to restrain progress indefinitely." As they wrote, "Biological models may not be able to substantially contribute to creating novel or enhanced pandemic-capable pathogens today," but "the essential ingredients to create highly concerning advanced biological models . . . may already exist or soon will. Establishment of effective governance systems now is warranted."

AI-Enabled Challenges to U.S. Biosecurity Measures

Fortunately, there are several significant **hurdles** to overcome when developing bioweapons. To start, there are many **millions, if not billions**, of viruses and bacteria in the world. Choosing the right one to weaponize matters immensely: Select a bacteria that multiplies **too slowly**, for instance, and the weapon will fail to disperse quickly enough, preventing maximum impact. One that mutates too fast may exhibit **different traits** than were intended by the time it is deployed. This is just the first of many complex steps in the pathway to develop a bioweapon, a simplified outline of which is illustrated in Figure 3.

Figure 3: Simplified Diagram of the Biological Weaponization Pathway


0. Avoid detection by law enforcement.
1. Select the organism.
2. Acquire the organism.
3. Multiply the organism.
4. Select the delivery method.
5. Package for weaponization.

Source: CSIS analysis of "Barriers to Biological Weapons Development: Potential Implications for Pathway Disruption," Countering WMD Journal 28 (Spring/Summer 2024).

Because no single intervention is sufficient to address the complexity, diversity, and evolving nature of biological threats, successfully mitigating biological risks from AI demands multiple **layers of defense** spread across multiple steps of the bioweaponization pathway. Different measures can offer unique choke points to developing bioweapons various stages of production; for example, effective AI-model security controls can prevent models from answering queries about planning and executing a biological attack (steps 0, 1, 3, 4, 5, and 6 in Figure 3). Likewise, nucleic acid synthesis screening can flag harmful sequence orders before they are produced (step 2 in Figure 3). These are just two biosecurity measures, but there are many more, such as physical security controls and management procedures, which are beyond the scope of this paper.

Still, some biosecurity measures may soon be inadequate for preventing AI-enabled biological threats as AI capabilities continue to improve. Indeed, evidence suggests that in certain instances, this may already be the case. The following section highlights key limitations in BDT security and synthesis-screening measures to date.

BDT SAFEGUARDS: CIRCUMVENTION RISKS IN THE EVO MODEL SERIES

Without appropriate safeguards, AI could help actors select (or create) the right viral sequences for bioweapons.[7] In anticipation of this risk, Evo 2's developers have **aimed to** ensure that the model does not contain information about pathogens that are infectious to humans, animals, and plants in its training dataset. This means Evo 2 **struggles** to predict the genetic sequences of certain harmful viruses, often gets what predictions it does make wrong, and effectively cannot simulate their genetic codes. What Evo 2's developers did is loosely analogous to teaching a student chemistry from a textbook with the chapters on explosives ripped out.

> *Without appropriate safeguards, AI could help actors select (or create) the right viral sequences for bioweapons.*

Still, if a future model like Evo 2 could indeed create the genetic sequences of entirely new organisms, there is no guarantee that such a model could not be modified to develop new pathogenic capabilities over time. Users with sufficient time and creativity could plausibly make novel, harmful organisms outside of the designated list that was removed from the model's training data. In keeping with the chemistry analogy, this would be like a student growing up to be such an overall expert in chemistry that they can eventually infer the fundamentals of explosives, even though they were not explicitly taught how to make them in school.

Another risk is that users could add the removed training data back into the model later. The results of the earlier Evo 1 model already indicate such a possibility. A 2024 academic **paper** on AI and biosecurity, which includes some of the same authors as the **Evo 2 technical paper**, stated the following:

> Creators of the Evo [1] model sought to guard against misuse by excluding "viral genomes that infect eukaryotic hosts" [including humans, plants, and animals] from the model's training data. They then published the model weights, as is currently the norm in academia. **Within weeks of Evo's release, other scientists had refined the model's published weights with data on viruses that infect humans.** The data in that case involved a benign virus family, but the case highlights that fine-tuning an open model is often much cheaper than training a new large model, and so oversight policies will need to account for the possibility of postrelease fine tuning. [emphasis added]

In short, eukaryotic viruses were also removed from Evo 1's pretraining dataset, yet users were able to add human viruses back into the model in a matter of weeks.

The Evo 2 **technical paper** did little to reassure that this could not happen again. The paper did not explicitly address how the developers sought to improve their security methodologies with Evo 2. All it said is that the authors

> collaborated with . . . multidisciplinary experts to reduce risks via data exclusion measures, safety and security evaluations, and population bias evaluations. By excluding genomic sequences of viruses that infect eukaryotes from our training data, we aimed to ensure our openly shared model did not disseminate the capability to manipulate and design pathogenic human viruses. **Task-specific post-training may circumvent this risk mitigation measure and should be approached with caution.**

The Evo 2 paper thus outlined a very similar biosecurity methodology to what was taken in developing Evo 1 (i.e., removing eukaryotic viruses from the pretraining dataset) and arrived at the same conclusion that these measures could be circumvented. As an **open-source, all-access tool**, the probability that this will happen again seems relatively high. At least for now, it appears such a weaponization pathway would nevertheless require a nontrivial amount of expertise to pull off. Still, as the Evo 1 experience highlights, experts could add harmful viral sequences back to the model, thus opening the door for less-sophisticated malicious actors.

Concerningly, the Evo model series represents two of only a small number of BDTs that possess any safeguards at all. An Epoch AI **survey** of 370 biological models—based on these tools' publications describing the developers' methodologies and processes—found that **fewer than 3 percent include safeguards** to mitigate risk, such as evaluations, access controls, or filtering and exclusion of training data. While many smaller, application-specific BDTs do not necessarily need the same levels of controls as more powerful models such as Evo 2, 3 percent is still an incredibly small number. If the Evo models represent some of the strongest biosecurity measures in the field—simply by having any safeguards at all—yet remain easily circumventable, this suggests a fundamental gap in how the wider BDT ecosystem approaches biological risk mitigation.

NUCLEIC ACID SYNTHESIS SCREENING: LIMITATIONS TO A LIST-BASED APPROACH

Luckily, there are **other hurdles** at various stages of bioweapons development beyond generating a suitably infectious and lethal viral sequence. For one, even if a BDT such as Evo 2 could simulate infectious organisms, a malicious actor would still need to acquire the physical strands of DNA (step 2 in Figure 3).

Unfortunately (from a security standpoint), some of the barriers are falling here, too. With technological advancements in **synthetic biology**, commercial firms can now "print" and ship custom-made synthetic nucleic acid (DNA or RNA) sequences around the world. As one synthesis company **advertises**, "If you know what you want, we will build it." Customers can place orders **over the internet** and receive their synthetic nucleic acid sample for a few cents per base pair (the fundamental unit of nucleic acids) within **two-to-four** business days. Thus, the previously mentioned viral horsepox DNA was acquired for as little as **$100,000** in 2017.

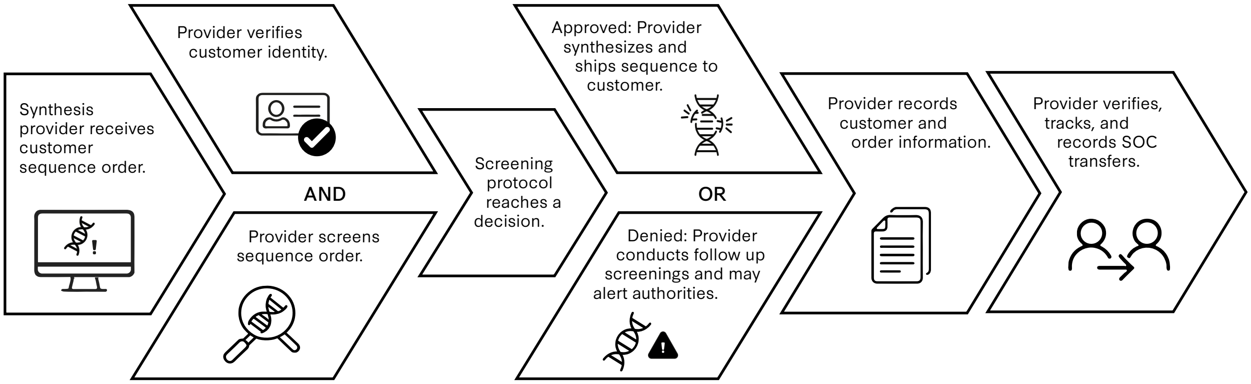
The U.S. government has issued guidelines for DNA synthesis firms to mitigate biosecurity risks from this **rapidly growing** industry. As of the time of writing, the Department of Health and Human Services (HHS) and the OSTP have published three major frameworks since 2010 to encourage companies to screen synthesis orders for harmful genetic content.[8] The 2023 HHS Framework, the most robust set of government guidelines to date, is outlined in bullet points below and illustrated in Figure 4. All three HHS and OSTP Frameworks, as well as notes on the Trump administration's May 5, 2025, **executive order** directing OSTP and relevant agencies to revise or replace the 2024 OSTP Framework (also **outlined** in the AI Action Plan), are summarized in Appendix B and Appendix C, respectively.

2023 HHS SCREENING FRAMEWORK GUIDANCE

In 2023, HHS updated its earlier 2010 guidance to reflect advancements in biotechnology capabilities, the growing synthesis industry, and an expanding set of associated biosecurity risks. This resulted in the **Screening Framework Guidance for Providers and Users of Synthetic Nucleic Acids**, the most widely referenced U.S. government framework for synthesis screening to date. The 2023 HHS framework issues recommendations for actors across the synthesis supply chain:

- **Providers** should screen all nucleic acid (i.e., single and double-stranded DNA and RNA) sequences for "best matches" to federally regulated agents, except when the sequence is also found in an unregulated organism or toxin. Moreover, "as soon as it is practical to do so," they should screen for nonregulated agents known to boost pathogenicity or toxicity in an organism. These sequences of concern (SOCs) are discussed in more detail below.
- **Providers and customers** should verify, track, and record any transfer of SOCs beyond the intended end user.
- **Providers and equipment manufacturers** should strengthen cyber and informational security measures to protect intellectual property and customer information.
- **Equipment manufacturers** should perform customer-screening processes, integrate SOC screening capabilities into equipment, and take security measures to ensure equipment cannot be overridden by unauthorized actors.

Figure 4: Simplified Step-by-Step Illustration of the 2023 HHS Screening Framework for Nucleic Acid Synthesis Providers


Synthesis provider receives customer sequence order.
Provider verifies customer identity.
AND
Provider screens sequence order.
Screening protocol reaches a decision.
Approved: Provider synthesizes and ships sequence to customer.
OR
Denied: Provider conducts follow up screenings and may alert authorities.
Provider records customer and order information.
Provider verifies, tracks, and records SOC transfers.

Source: Department of Health and Human Services, National Security Commission on Emerging Biotechnology, and CSIS analysis.

While these frameworks are a commendable first step, they could soon become obsolete due to the increasing convergence of AI and biotechnology. Imagine the following scenario in the not-too-distant future: A DNA synthesis company receives a request from a customer asking it to produce a sequence. Following the latest government guidelines, a company employee dutifully screens the sequence against the company's database of SOCs, including federally regulated agents and toxins. Nothing is flagged. However, it is plausible that the customer used an AI tool to develop this sequence—these tools are increasingly popular among scientists, and the employee has heard of other recent instances in which SOC databases did not catch harmful AI-generated sequences. Still, government guidelines technically indicate that nothing is wrong. How should the employee proceed?

This scenario may seem dramatic, but it is already frighteningly close to reality. In December 2024, experts representing several major biotechnology companies and biosecurity organizations published a **study** in which they stated that widely available AI protein models are "now making it possible to generate . . . proteins potentially hundreds of mutations away from the closest known natural protein." To test whether these AI-generated proteins could circumnavigate synthesis companies' screening tools, the authors generated tens of thousands of harmful sequences that were functionally equivalent to but sequentially different from known SOCs. They then fed these sequences through biosecurity screening software (BSS) used by commercial providers in a secure and confidential environment.

The authors **reported** that several companies' software, "including screening methods in use at major nucleic acid suppliers, could not reliably detect such AI-reformulated toxins and viral proteins," with "up to 100% of variants from certain proteins passing undetected through at least one BSS." By adjusting the screening software, including integrating a model that can recognize signs of AI-generated protein variants, the authors were ultimately able to significantly boost detection rates up to 97 percent. Still, the study concluded that "in the long run, we should not expect sequence-based [biosecurity] strategies alone—like those tested and developed in this study—will be sufficient, as we envision a future in which AI-assisted generation of proteins produces sequences unlike any seen in nature."

Indeed, both this study as well as the future scenario provided above serve to highlight three **key limitations** with current U.S. government screening guidelines. These are outlined briefly below and described in greater detail in Appendix D.

1. **The current list of 63 regulated agents and toxins will likely fail to detect the possibly infinite new or modified pathogens designed by future BDTs.** To date, there are only 63 regulated toxins, pathogens, and bacteria on the **Select Agents and Toxins List** compiled by HHS and Department of Agriculture (USDA). When including export-controlled items on the Bureau of Industry and Security's **Commerce Control List** for international orders, there are only approximately 150 unique items.[9]

   In theory, the frameworks' screening mechanisms work until an AI model creates the 64th (or, for international orders, say the 151st) deadly organism that is not yet on one of these lists. There are thus serious risks that these lists of regulated sequences are an increasingly **inadequate** line of defense, especially when just one sufficiently different genomic sequence could be enough to unleash a highly contagious virus or even start the next global pandemic.

   Indeed, an **April 2025 report** commissioned by Congress from the National Security Commission on Emerging Biotechnology (NSCEB) **called** current U.S. biosecurity measures like the regulated agents lists "blunt and reactive" for this reason. A **2018 biodefense report** from NASEM went further, arguing that this "overreliance on the Select Agent list is a systemic weakness affecting many aspects of the United States' current biodefense mitigation capability." Even the 2023 HHS framework **recognized** this shortcoming, acknowledging that advancements in biotechnology may assist users with "engineering pathogenic or toxic proteins with completely novel sequences . . . that are not a match to any known sequence." Concerningly, the framework leaves it up to providers to develop best practices to address such issues themselves.

2. **Efforts to expand synthesis-screening protocols to capture additional sequences of concern are nascent and could quickly fall behind AI-generated sequence capabilities.** Among the largest updates to U.S. government guidelines on synthesis screening in recent years has been to encourage providers to expand the scope of their screening processes to include additional harmful, nonregulated genetic sequences. While this is commendable for improving biosecurity in theory, there are several key challenges to implementing it in practice:

   - There is currently **no consensus** among industry and government stakeholders regarding what constitutes an SOC.
   - There is **no standardized database** of SOCs beyond federally regulated lists.
   - The generative capabilities of future AI biology models threaten to make new or existing SOC databases continuously incomplete.

3. **Current government frameworks encourage providers and other actors to develop their own security best practices, increasing the likelihood of a fragmented biosecurity landscape in which companies who skimp on safety hold a competitive cost or speed advantage over those who implement more rigorous measures.** Despite iterating security best practices for over a decade, U.S. government guidelines for DNA synthesis companies remain largely high level. While high-level guidance arguably increases flexibility and reduces

the regulatory burden for large companies, it may leave smaller companies to develop their own protocols from scratch. Thus, these guidelines serve as stopgap measures at best, rather than detailed instructions for synthesis companies.

Patchwork efforts may also inadvertently punish good actors that adopt stricter security measures. Companies report high costs associated with adhering to current government screening guidelines, which are **expected** to become an increasingly **large** part of synthesis companies' total overhead costs. In a worst-case scenario, high screening costs would give companies commercial incentives to cut corners on safety best practices. This same dynamic could play out at a global scale: Customers could exploit the current lack of international synthesis security standards and choose foreign providers that employ more relaxed measures, if any, to produce dangerous sequences abroad. As the Covid-19 pandemic showed, highly contagious pathogens from anywhere can quickly become a health crisis everywhere.

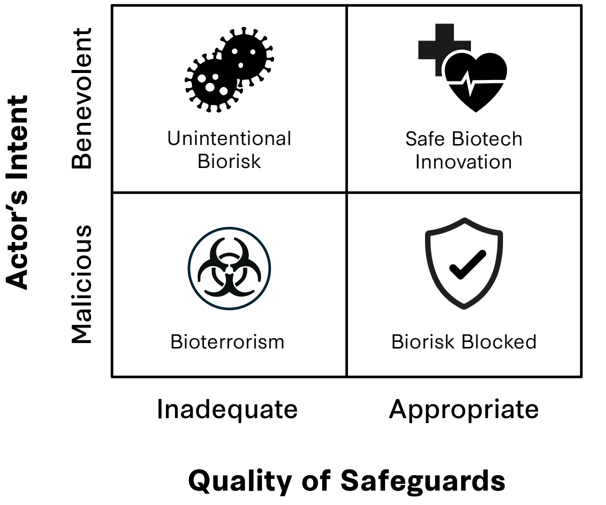
Recommendations

The **AI Action Plan** rightly identifies U.S. government investment into biosecurity as an essential component to the United States' national security strategy. In practice, however, treading the line between mitigating viral sequences used to cause catastrophe while enabling those used for life-saving research is undoubtedly challenging.

U.S. policymakers have the tough job of ensuring biosecurity measures are stringent enough that bad actors cannot develop bioweapons, but not so stringent that they limit good actors from developing genetic sequences–including pathogenic ones–for beneficial scientific efforts such as drug discovery or vaccine development (see Figure 5). Lawmakers also have the dual responsibility of strengthening biosecurity practices without adding undue burden to companies working in the AI, DNA synthesis, or other sectors.

To avoid undermining U.S. competitiveness in biotechnology, policymakers should at least enhance security measures that are no more burdensome than existing requirements (and are ideally less burdensome). They should also carefully consider the potential opportunity costs of constraining frontier AI and biotechnology innovation with stricter security measures. This requires developing targeted policy interventions that do not stifle commercial AI and biotechnology progress or unduly **disrupt** academic autonomy. Finally, given the inherently global nature of biorisk and biosecurity, robust mitigation measures demand international engagement with a wide variety of stakeholders from lower-, middle-, and high-income countries alike.

Figure 5: Simplified Biosecurity Policy Matrix



Source: CSIS analysis.

If developed and implemented thoughtfully, clear and proportionate safeguards could accelerate biotechnology innovation and adoption by building trust and helping seize the incredible promises that AI holds for the scientific and biomedical fields. What follows are three recommendations the authors believe would strengthen biosecurity in the age of AI without slowing down the U.S. AI or biotechnology industry.

1. **U.S. lawmakers should fund NIST and CAISI to continue and expand their critical AI and biosecurity work.**

U.S. government guidelines will require continuous revision and updating as AI increasingly poses novel risks to biosecurity, including but not limited to the threats assessed in this paper. Moreover, the scope of AI-enabled biosecurity risks will shift over time as new models develop new capabilities.

This reality demands government expertise, capacity, and industry engagement that the National Institute of Standards and Technology and, within it, the recently announced **U.S. Center for AI Standards and Innovation** (formerly the **AI Safety Institute**) are well-positioned to meet. These institutions are doing unique and urgently needed work to prevent the misuse of AI and synthetic biology for deliberate and accidental biological attacks. Ongoing biosecurity-related efforts at NIST and CAISI include:

NIST:

- developing high-level requirements for SOC databases, including standardize metadata and database access requirements, in partnership with the Engineering Biology Research Consortium (EBRC) and other relevant stakeholders;
- developing test datasets and benchmarking tools to assess and maintain the performance of nucleic acid synthesis tools;
- working to harmonize global nucleic acid sequence-screening practices within the International Organization for Standardization ISO and other standards-development organizations; and
- advancing AI capabilities to enable the prediction of how sequences will function (specifically developing standard methods for the collection of large, high-quality datasets for training and testing AI tools).

CAISI:

- conducting pre-deployment evaluations of U.S. and foreign commercial AI models for capabilities and demonstrable risks to national security, including biological risks;
- working with NIST organizations to develop guidelines, best practices, and voluntary standards for measuring and improving the security of AI systems; and
- coordinating with other federal agencies and authorities, including the Department of Defense, the Department of Energy, and the Office of Science and Technology Policy, to conduct AI system evaluations and assessments.

NIST and CAISI have already received clear federal mandates to lead critical AI and biosecurity initiatives. Most recently, the **AI Action Plan** identifies NIST and CAISI as primary contacts within the U.S. government to lead security testing of frontier AI models for biological and other national security risks, demonstrating continuity with the Biden administration's 2024 **national security memorandum** (NSM) on AI. However, the AI Action Plan also expands the scope of NIST and CAISI's responsibilities beyond those listed in the NSM, such as helping to monitor foreign frontier AI developments with potential U.S. national security implications and developing 'financial markets' for compute.

Yet just as the Trump administration is tasking NIST and CAISI with deeper and broader responsibilities under the Action Plan, it is also proposing significant reductions in their budgets. President Trump's fiscal year 2026 budget recommendations to Congress proposed a **$325 million cut** to NIST's funding–a nearly **30 percent** reduction from its FY 2025 budget of around **$1 billion**–citing the agency's longstanding support of the "radical climate agenda." As part of this proposal, scientific and technical research and services would see a nearly **17 percent** cut compared to the FY 2025 budget, as well as a **reported** elimination of nearly 500 full-time employees and a further $10 million in pay cuts to remaining staff.

For now, the specific NIST offices that stand to be the most impacted by these proposed cuts are unclear. Given the importance of CAISI to the AI Action Plan, it is likely that the center will be shielded from the worst effects of the proposed FY 2026 budget cuts. Indeed, the plan recommends that the government prioritize *recruiting* top AI research talent to federal agencies including NIST and CAISI, rather than cutting staff.

But recruiting and retaining top AI talent requires significant investment from the U.S. government, not budget cuts and layoffs. Highly competitive industry salaries already make attracting AI researchers to government service challenging, and executing the expanded scope of NIST and CAISI's responsibilities under the AI Action Plan will demand more talent and funding, not less.

Moreover, shrinking NIST's budget and expertise would further degrade the U.S. government's ability to prepare for and respond to biorisks just as the **FY 2026 budget proposal** threatens enormous cuts to other health, safety, and biosecurity-related agencies, including the National Institutes of Health (–$18 billion), the Centers for Disease Control and Prevention (–$3.6 billion), and the Administration for Strategic Preparedness and Response (–$240 million). It would also coincide with increasingly dwindling staff across key federal biosecurity-related agencies and a leaderless White House Office of Pandemic Preparedness and Response Policy following the recent **resignation** of top official Gerald Parker.

If the Trump administration is as serious about investing in biosecurity as AI Action Plan claims, it should not underfund NIST and by extension CAISI - two of its greatest technical resources. At a minimum, U.S. policymakers and lawmakers should reconsider the proposed 30 percent reduction in NIST funding, especially given its proposed cuts to other relevant agencies responsible for population health and biosecurity. Better yet, lawmakers should look to authorize additional funding for NIST (and by extension CAISI) given the administration's goals to recruit talent and expand the institutions' evaluation and research work.

> *If the Trump administration is as serious about investing in biosecurity as AI Action Plan claims, it should not underfund NIST and by extension CAISI - two of its greatest technical resources.*

2. **CAISI should leverage the interagency Testing Risks of AI for National Security (TRAINS) Taskforce and the international network of AI Safety Institutes to develop evaluations of frontier BDTs for high-impact capabilities.**

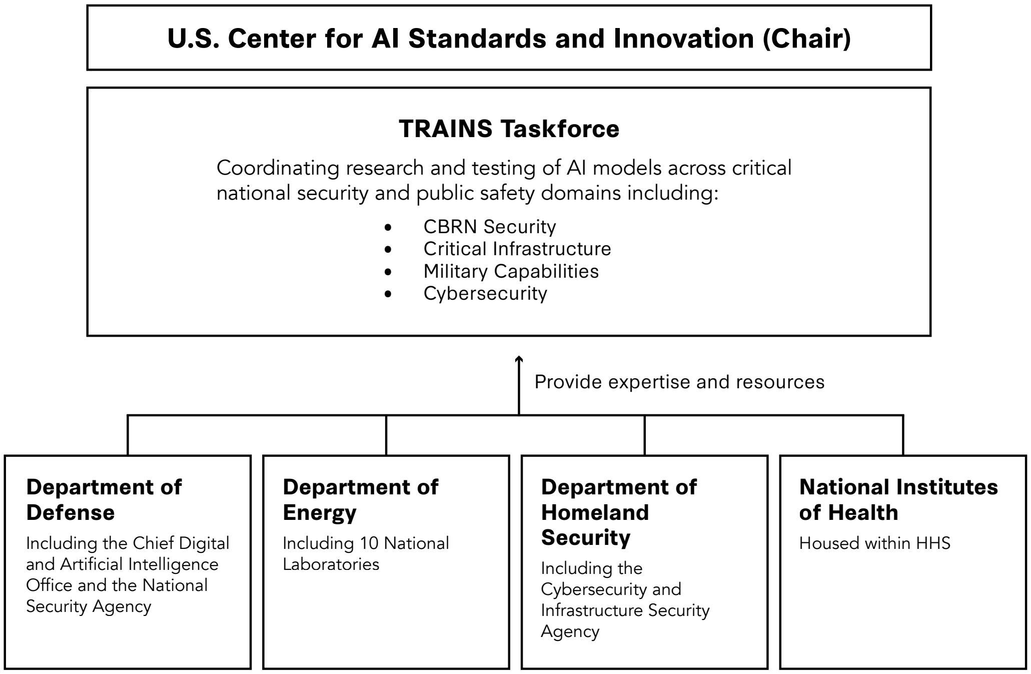
Advancements in BDTs pose unique risks to U.S. biosecurity that should be monitored carefully as they evolve (see the section of this paper on BDTs above). CAISI should conduct BDT model evaluations to monitor frontier BDTs for emerging high-impact biological capabilities, including but not limited to novel pathogen design and modifying existing pathogens to evade screening procedures. Such evaluations would directly support CAISI's stated **goal** of mitigating demonstrable national security risks, including biosecurity threats, as well as the AI Action Plan's **recommendation** that the federal government develop national security-related AI evaluations.

To avoid stifling innovation, only the **most powerful** BDTs should require testing, defined by proxy metrics such as training-compute thresholds and training-dataset size.[10] To capture the full breadth of the frontier BDT landscape, testing should encompass commercial frontier models, models developed by noncommercial organizations (including universities and nonprofits as in the case of Evo 2), and models released under noncommercial licenses (such as **ESM3-open**).

To do so, CAISI may need to expand its recently narrowed **focus** on commercial AI systems. It is currently unclear whether CAISI's definition of commercial systems exclusively covers AI systems developed by companies, such as LLMs produced by OpenAI and Anthropic, or nonprofit-developed and noncommercially licensed systems as well. To capture the broadest array of high-impact frontier AI systems, CAISI should employ this expanded definition of commercial systems in its evaluations and research on frontier biological models.

CAISI need not develop BDT evaluation capabilities alone. Rather, it should look to other relevant agencies and existing resources such as the interagency TRAINS Taskforce, dedicated to conducting research and testing of AI models across national security and public safety domains, including for biosecurity risks (see Figure 6). The TRAINS Taskforce was already chaired by the former U.S. AI Safety Institute, offering the opportunity for continuity with the new CAISI. Furthermore, it brings in other agencies with important expertise and resources relevant to AI biological evaluations, including the Department of Energy and the National Institutes of Health.

Figure 6: Organizational Chart of the TRAINS Taskforce


U.S. Center for AI Standards and Innovation (Chair)
TRAINS Taskforce
Coordinating research and testing of AI models across critical national security and public safety domains including:
• CBRN Security
• Critical Infrastructure
• Military Capabilities
• Cybersecurity
Provide expertise and resources
Department of Defense
Including the Chief Digital and Artificial Intelligence Office and the National Security Agency
Department of Energy
Including 10 National Laboratories
Department of Homeland Security
Including the Cybersecurity and Infrastructure Security Agency
National Institutes of Health
Housed within HHS

Source: NIST and CSIS analysis.

Internationally, CAISI could tap into the **global network** of AI Safety Institutes for additional capacity and expertise from trusted partners abroad. While some BDTs developed by U.S. entities may be restricted to testing within the United States due to security and intellectual property reasons, there are nevertheless many open-source models that can be effectively researched and evaluated for biosecurity risks elsewhere. International cooperation and information sharing on this issue could also help enhance the United States' understanding of foreign AI biology capabilities and harmonize model-evaluation metrics important for setting global standards.

3. **OSTP and other relevant agencies should develop a plan to integrate AI capabilities into a standardized synthesis-screening system for U.S. synthesis providers.**

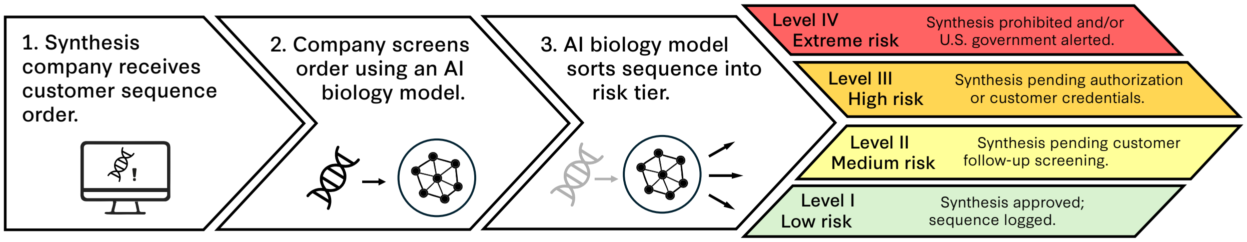
Thus far, this paper has highlighted a future in which AI biology models with capabilities like Evo 2's could soon help malicious actors develop entirely new categories of bioweapons. However, such models also present an exciting opportunity for strengthening U.S. biodefense.

A future AI biology tool with advanced predictive capabilities to determine sequence structures, functions, and mutational effects could be applied to identify and model potential pathogens. DNA synthesis companies could leverage these capabilities in their screening processes to simulate whether variations in genetic sequences would increase pathogenic or toxic traits, how a novel genetic sequence would behave, and to flag new sequences of concern before synthesis even occurs.

Using a future AI biology model with improved predictive capabilities like Evo 2's, synthesis companies could screen all sequence orders to determine whether it does the following

- matches a federally regulated sequence;
- matches a known SOC outside of this list; and/or
- would imply novel pathogenic or toxic functions that could constitute a new SOC within a certain confidence interval (CI).

Figure 7: Simplified Illustration of a Hypothetical AI-Enabled Nucleic Acid Screening Process Using an AI Biology Model



1. Synthesis company receives customer sequence order.
2. Company screens order using an AI biology model.
3. AI biology model sorts sequence into risk tier.
Level IV Extreme risk
Synthesis prohibited and/or U.S. government alerted.
Level III High risk
Synthesis pending authorization or customer credentials.
Level II Medium risk
Synthesis pending customer follow-up screening.
Level I Low risk
Synthesis approved; sequence logged.

Source: CSIS analysis.

The model could then take appropriate next steps according to the sequence's biosecurity risk following a **tiered system** (see Figure 7):

- **Low Risk:** Known or novel sequences with no structural similarity to known pathogens or toxins. Sequences would be logged but cleared for synthesis without delay.
- **Medium Risk:** Sequences with some structural similarities to known agents, such as attenuated sequences used in vaccines, but below critical thresholds. These might trigger a customer screening follow-up, such as confirming researcher credentials or requesting further project information.
- **High Risk:** Sequences that closely resemble known pathogens or that the AI predicts would exhibit harmful properties within a certain confidence interval (CI). These sequences could require explicit approval from a government authority or the customer to provide credentials before synthesis.
- **Extreme Risk:** Known or novel sequences confidently predicted (e.g., at an 80 percent CI or higher) to have pathogenic/toxic characteristics of a pandemic or similarly extreme scale. This category would halt synthesis and alert U.S. government authorities of the sequence order.

Additional considerations for a government-led AI-enabled synthesis-screening system, including security issues, outstanding challenges, and an example of an earlier government effort to integrate AI into synthesis screening–such as the Functional Genomic and Computational Assessment of Threats (Fun GCAT) run by the Intelligence Advanced Research Projects Activity (IARPA)–are outlined in Appendix D.

Conclusion

AI promises to revolutionize the field of biology and bring enormous benefits to humanity. But rapid advancements in LLMs and BDTs also pose new and emerging biosecurity threats, and current U.S. biosecurity measures are ill-equipped to meet these challenges. Experts warn that critical safeguards embedded within BDT models are already circumventable post-deployment, suggesting grim implications as model capabilities improve, and U.S. government guidelines for list-based synthesis screening measures remain blunt tools that will likely fail to detect future AI-generated sequences that do not match known agents and toxins.

The Trump administration has made promising first steps by addressing these challenges in its AI Action Plan and May 2025 **executive order**. However, the real test will be how quickly it is able to translate its high-level recommendations into actionable policies that keep pace with technological breakthroughs. U.S. policymakers also face the undoubtedly tough job of preventing biothreats while also minimizing opportunity costs by adding further burdens to industry.

The recommendations outlined in this paper will help the United States to pursue such a high-return, low-cost biodefense strategy. As the barriers to bioterrorism continue to fall in the age of AI, U.S. policymakers must strive to ensure that developing bioweapons remains as difficult and costly as possible. The actions they take may determine whether bioterrorism remains the realm of a few evil geniuses, or whether it falls within reach of evil morons as well. ■

About the Authors

***Georgia Adamson** is a former research associate at the Wadhwani AI Center at the Center for Strategic and International Studies (CSIS) in Washington, D.C. **Gregory C. Allen** is a senior adviser to the Wadhwani AI Center at CSIS.*

Acknowledgements

The authors would like to thank J. Stephen Morrison, Michaela Simoneau, and Sophia Hirshfield from the CSIS Global Health Policy Center for lending their support and expertise to this project.
The authors would also like to thank Andrew Hebbeler, Shivshankar Sundaram, Melissa Hopkins, Jassi Pannu, Jaime Yassif, and Jake Beal for their helpful insights and/or feedback on earlier drafts of this report.

The authors would also like to thank Epoch AI for sharing its dataset on existing biological model safeguards.

The views and any mistakes made in this paper are the authors' own.

This report is made possible by general support to CSIS. No direct sponsorship contributed to this report.

This report is produced by the Center for Strategic and International Studies (CSIS), a private, tax-exempt institution focusing on international public policy issues. Its research is nonpartisan and nonproprietary. CSIS does not take specific policy positions. Accordingly, all views, positions, and conclusions expressed in this publication should be understood to be solely those of the author(s).

Cover Photo: Chom via Adobe Stock

Appendix A

Advancements in the Evo Foundation Model Series (2024-25)

Between February 2024 and February 2025, the Evo foundation model series demonstrated significant improvements in BDT capabilities. Figure A-1 below compares Evo 1 and Evo 2 model specifications, such as size and sequence context length. Figure A-2 compares model performance capabilities, including sequence prediction, generation, and design.

Figure A-1: Comparison of Evo 1 and Evo 2 Model Specifications

Model Specifications	Evo 1	Evo 2
Release date	February 27, 2024	February 19, 2025
Tokens	c. 340 billion (single-cell genomes)	> 9.3 trillion (nucleotides)
Parameters	7 billion	7 billion & 40 billion
Training data	Bacterial and phage genomes	Bacterial, achaeal, phage, and eukaryotic genomes
Training hardware	64 Nvidia H100 GPUs & 128 Nvidia A100 GPUs	2,000 Nvidia H100 GPUs
Training compute estimates (FLOPS)	1.26 x 10^{22}	2.25 x 10^{24}
Sequence context length	131,000 base pairs	1 million base pairs

Source: Arc Institute (Evo); Arc Institute (Evo 2); "Sequence Modeling and Design from Molecular to Genome Scale with Evo," Science 386, no. 6723 (November 15, 2024); and bioRxiv.

Figure A-2: Sample Performance Comparisons Between Evo 1 and Evo 2

Model Performance Metrics	Evo 1	Evo 2
Predictive capabilities	Limited to bacteria and phages	State-of-the-art performance on bacteria, phages, and eukaryotes (ex-human viruses)
Generative capabilities at genome scale	Smaller sequences approximately the size of the smallest "minimal" bacterial genomes	Larger sequences at the scale of whole human mitochondrial genomes, minimal bacterial genomes, and entire yeast chromosomes
Design capabilities	Limited design capabilities able to generate specific CRISPR-Cas molecular complexes	Complex design capabilities able to generate new genomic sequences using inference-time scaling to guide generation towards desired properties

Source: Sequence Modeling and Design from Molecular to Genome Scale with Evo," Science 386, no. 6723 (November 15, 2024); and bioRxiv.

Appendix B

Overview of U.S. Government Guidelines for Nucleic Acid Synthesis Screening to Date

As of the time of writing, the U.S. government has issued three major frameworks to encourage synthesis providers to screen incoming orders for harmful genetic content. These frameworks are outlined in high-level bullet points below. Comments on the Trump administration's upcoming framework can be found in Appendix C.

2010: HHS Screening Framework Guidance for Providers of Synthetic Double-Stranded DNA

In 2010, HHS issued the first comprehensive set of government guidelines to establish baseline safety and security best practices around double-stranded DNA technology. The voluntary framework encouraged synthesis companies (which they call "providers") to:

- Verify customer identities.
- Screen all orders for sequences derived from or encoding regulated sequences listed in the HHS/USDA Select Agents and Toxins List and, for international orders, the Bureau of Industry and Security's CCL.
- Conduct follow-up screenings if a customer or sequence order raises concerns and contact U.S. authorities if appropriate.
- Maintain records of all screening protocols, customers, orders, and sequence information for a minimum of eight years.

2023 HHS Screening Framework Guidance for Providers and Users of Synthetic Nucleic Acids

In 2023, HHS updated the 2010 guidance to reflect advancements in biotechnology capabilities, the growing synthesis industry, and an expanding set of associated biosecurity risks. Updates encourage actors across the synthesis supply chain to undertake the following:

- **Providers:** Screen all nucleic acid (i.e., single and double-stranded DNA and RNA) sequences for "best matches" to federally regulated agents, except when the sequence is also found in an unregulated organism or toxin. Additionally, "as soon as it is practical to do so," screen for SOCs.
- **Providers and customers:** Verify, track, and record any transfer of SOCs beyond the intended end user.
- **Providers and equipment manufacturers:** Strengthen cyber and informational security measures to protect intellectual property and customer information.
- **Equipment manufacturers:** Perform customer screening processes, integrate SOC screening capabilities into equipment, and take security measures to ensure equipment cannot be overridden by unauthorized actors.

2024 OSTP Framework for Nucleic Acid Synthesis Screening

In October 2023, the Biden-Harris AI executive order (Executive Order 14110) instructed OSTP to establish a new framework for screening synthetic nucleic acid sequences in consultation with HHS and other federal agencies. The resulting framework largely continues earlier 2023 HHS guidance for screening customers and sequence orders, with a few additional updates for providers and manufacturers:

- **Providers:** Post a statement on the provider's website that it adheres to the screening framework and/or provide a copy to federally funded customers or federal agencies upon request.
- **Providers and equipment manufacturers:** Report potentially illegal orders detected in the screening process to appropriate U.S. authorities.
- **Providers and equipment manufacturers:** Enhance cyber and supply chain security in accordance with NIST's Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations.

Appendix C

Notes on the Trump Administration's Forthcoming Updates to the 2024 OSTP Framework for Nucleic Acid Synthesis Screening

On May 5, 2025, President Trump issued an executive order on Improving the Safety and Security of Biological Research with important updates to U.S. government guidelines on nucleic acid synthesis screening. The executive order directs the director of OSTP, in coordination with the assistant to the president for national security affairs (APNSA) and the heads of relevant agencies, to revise or replace the Biden administration's 2024 OSTP framework within 90 days. While most of the details about this update are not yet public, the executive order provides a preview of three notable changes to the existing government guidance:

- **Expansion of the framework to include non-federally funded entities:** Within 180 days of the executive order, the director of OSTP, in coordination with APNSA, the Office of Management and Budget, and other relevant agencies, will develop and implement an updated framework that includes non-federally funded entities. This expands the scope of the Biden administration's 2024 framework, which applied only to federally funded research entities (representing the majority of all synthesis orders), to also include the private sector. Any gaps in executive branch authorities to meet this objective will be addressed through a legislative proposal within the 180-day period.
- **Enforcement requirements:** The updated framework will also mandate enforcement provisions in every life science research contract or grant award. This includes terms requiring the contract signatory or grant recipient to:
  - Comply with the executive order and other applicable regulations.
  - Certify that they do not operate, participate in, or fund any dangerous gain-of-function research or other life science research in foreign countries that could cause significant societal consequences or generate unnecessary national security risks.
  - Acknowledge that a violation of the terms of this executive order or other applicable regulations may also constitute a violation by the signatory or recipient's employer or institution, potentially subjecting them to the immediate revocation of ongoing federal funding and a five-year bar from future federal life sciences grants.
- **Framework revision timelines:** Moving forward, OSTP and other relevant agencies will review and revise the updated framework at least every four years or as appropriate, instead of every two years as the 2024 framework states.

Note that the forthcoming framework will fulfil the Trump administration's first recommendation in the "Invest in Biosecurity" section of the AI Action Plan on this issue.

Appendix D

Limitations to Current HHS and OSTP Frameworks for Nucleic Acid Synthesis Screening (Expanded Version)

1. **The current list of 63 regulated agents and toxins will likely fail to detect the possibly infinite new pathogens made by BDTs.**

   To date, there are only 63 regulated toxins, pathogens, and bacteria on the HHS/USDA Select Agents and Toxins List. When including export-controlled items on the Bureau of Industry and Security's CCL for international orders, there are only roughly 150 unique items.

   In theory, the frameworks' screening mechanisms work until an AI model creates the 64th (or, for international orders, say the 150st) deadly organism that is not yet on one of these lists. There are thus serious risks that these lists of regulated sequences are an increasingly inadequate line of defense, especially when just one sufficiently different genomic sequence could be enough to start the next global pandemic.

   Indeed, an April 2025 report commissioned by Congress from the NSCEB calls current U.S. biosecurity measures like the regulated agents lists "blunt and reactive" for this reason. A 2018 biodefense report from NASEM goes further, arguing that this "overreliance on the Select Agent list is a systemic weakness affecting many aspects of the United States' current biodefense mitigation capability." Even the 2023 HHS framework recognizes this shortcoming, acknowledging that advancements in biotechnology may assist users with "engineering pathogenic or toxic proteins with completely novel sequences . . . that are not a match to any known sequence." Concerningly, the framework leaves it up to providers to develop best practices to address such issues themselves.

2. **Efforts to expand synthesis-screening protocols to capture additional SOCs are nascent and could also quickly fall behind AI-generated sequences.**

   Among the biggest updates to U.S. government guidelines on synthesis screening in recent years has been to encourage providers to expand the scope of their screening processes beyond the Select Agents and Toxins List and CCL. In 2023, the updated HHS framework acknowledged that current lists of federally regulated agents

   > do not represent the entirety of potential risks to public health . . . that could arise from the misuse of synthetic nucleic acids. Non-regulated pathogens and toxins, as well as other novel types of nucleic acid sequences, may also pose significant risks if they are misused. To minimize these risks, a shift is needed from relying solely on lists of regulated pathogens and toxins to also assessing the risks associated with other nucleic acid sequences that may contribute to pathogenicity or harm if introduced into new genetic frameworks.

   The 2023 framework adds that the expanded definition of SOCs should take effect "as soon as it is practical to do so," while the subsequent OSTP framework set the specific deadline of **October 13, 2026**.[11]

Encouraging providers to cast a wider net in their screening protocols marks a shift from a taxonomy-based approach (whether a sequence specifically matches the biological classification of a regulated organism) to a risk-based one (whether it can cause harmful effects writ large). While this may be a good idea for improving biosecurity in theory, there are several key challenges to implementing it in practice:

2a. There is currently no consensus among industry and government stakeholders regarding what constitutes an SOC. Even U.S. government guidelines disagree: The **2023** HHS framework states that an SOC includes sequences that "may contribute to pathogenicity or harm if introduced into new genetic frameworks," whereas the 2024 OSTP version states that an SOC covers sequences that are concretely "known to contribute to pathogenicity or toxicity." These differences in definition may seem small but could matter a lot when determining which sequences do or do not get screened and how companies approach their safety mitigation requirements.

2b. There is no standardized database of SOCs beyond regulated lists. Indeed, the **2023 HHS framework** recognizes that such a database "may not yet exist" and encourages providers to develop their own. Fortunately, there are already several search tools and databases for screening sequence orders. For example, the HHS National Center for Biotechnology Information's (NCBI's) **freely accessible** Basic Local Alignment Search Tool (**BLAST**) compares sequences to NCBI **genetic databases** (including viral ones) and determines the statistical significance between matches. Similarly, the commercial screening platform **ThreatSEQ** checks DNA orders against an SOC database it claims covers 100 percent of regulated agents and "virtually all known bacterial human/zoonotic pathogens."

Individually, these are valuable screening resources. However, the multiplicity of these datasets also risks creating a patchwork of variable screening methods depending on the different assumptions and sequences each includes. A January 2025 report from the EBRC notes the following:

> Different screening philosophies, threat models, and perspectives can greatly influence what is or is not included in a SOC database, even when a database developer is trying to only include biosecurity relevant sequences from regulated organisms. For example, amongst existing screening tools, some assume a sequence from a regulated agent is "innocent until proven guilty." Other tools assume these sequences are "guilty until proven innocent," which would greatly expand both the number and types of sequences of concern in a database.

The report concludes that "screening would likely be more consistent across Providers if their screening systems referenced the same SOC database(s)." Similarly, a **2020 HHS survey** states that most in industry find the HHS BLAST adequate for screening, but that "the lack of a definitive database of biothreat sequences was identified as a gap" in U.S. biosecurity.[12]

2c. The generative capabilities of BDTs threaten to make new or existing SOC databases continuously incomplete. Currently, SOC databases rely upon published scientific research to determine whether a sequence should be included in the dataset. This is a time- and research-intensive approach that can create multiple gaps in the database and contribute to a **static infrastructure** for pathogen detection. According to the February 2025 **EBRC report**:

> It is not uncommon for developers of SOC databases to find previously published papers that cause them to add a sequence or change its annotation [in a database]. . . . New research is constantly published that may impact whether a given sequence is considered "of concern."

Thus, the report concludes that developing a comprehensive database of SOCs would

> require *significant* resources to comprehensively reflect all SOCs in the published literature. . . . Even if a database developer was able to accomplish such a feat, the database would quickly become outdated without significant investment in its maintenance.

AI may soon compound this issue, potentially generating novel pathogens faster than scientists can publish articles. Compared to typical timelines in AI, publishing in scientific journals occurs at a **snail's pace**. By the time scientists have studied a new AI-generated viral sequence, then published a paper that may (or may not) be incorporated into an SOC database, it could be too late. Imagine, for instance, a scenario in which synthesis companies manufacture an AI-generated viral sequence simply because a scientific paper is held up in a lengthy peer review process. This assumes that scientists are even aware of the novel sequence—most malicious actors are unlikely to write academic papers on their plans, at least not without executing them first.

3. **Current government frameworks encourage providers and other actors to develop their own security best practices, increasing the likelihood of a fragmented biosecurity landscape in which companies who skimp on safety hold a competitive advantage over those who implement more rigorous measures.**

Despite iterating security best practices for over a decade, U.S. government guidelines for DNA synthesis companies remain largely high level. Of the three frameworks published as of the time of writing, only the 2010 HHS document was longer than 14 pages. These frameworks are so short principally because they place **substantial responsibility** on providers, manufacturers, and other actors to develop their own best practices.

While high-level guidance arguably increases flexibility and reduces the regulatory burden for large companies, it may leave smaller companies to develop their own protocols from scratch. Thus, these guidelines serve at best as stopgap measures rather than detailed instructions for synthesis companies.

Fortunately, several ongoing initiatives are working to coordinate and standardize synthesis-screening best practices. For example, the **International Gene Synthesis Consortium**, an industry-led organization, proactively shares processes and screens orders against a communal database of internationally regulated pathogens and toxins. According to the **consortium**, it represents "approximately 80% of gene synthesis capacity worldwide," with members including multinational biotech firms such as GenScript as well as smaller synthesis labs. Similarly, the **International Biosecurity and Biosafety Initiative for Science** and the **Bio Funders Compact**, led by the Nuclear Threat Initiative and the Coalition for Epidemic Preparedness Innovations, are working to integrate biosecurity best practices into the bioscience research and development lifecycle. Furthermore, NIST is currently working with international standards-development organizations and the EBRC to develop global best practices for screening protocols and tools (see the recommendations section of this paper).

Efforts to coordinate synthesis safety protocols are commendable, but they also demonstrate the incompleteness of government efforts to date. Leaving companies to develop their own best practices could encourage a patchwork of security measures and ultimately a fragmented biosecurity landscape, in which malicious actors could shop around for providers with the laxest safeguards and industry would generate no meaningful economies of scale in implementing screening methodologies.

Indeed, patchwork efforts could inadvertently punish good actors that adopt stricter security measures. In 2024, HHS officials published a **review** of industry and other stakeholder feedback, which informed the updated 2023 HHS framework, reporting:

> According to respondents, **implementing the 2010 screening framework is expensive, costing ~$15/order, and this cost has remained flat over the years whereas the cost of synthesis has decreased. Cost is reportedly driven by the need for a PhD in Bioinformatics to interpret sequence hits. The burden of screening may be expected to increase as the size of databases against which to match BLAST sequences is increasing (i.e., as the definition of SOCs expands beyond solely sequences from regulated pathogens and toxins).** Providers indicated that they need new annotated data resources, tools, and approaches to keep biosecurity from becoming a leading component of the per-bp [base pair] cost, given that nucleic acid synthesis costs have consistently decreased over time.

Currently, 15 dollars per order seems a small price for safety given that sequence orders typically **start** at tens of thousands of dollars. However, as the DNA synthesis industry expands and the cost of synthesis **falls**, the fixed price of adhering to government frameworks could represent an increasingly **large** part of synthesis companies' total overhead costs. Companies **report** that current screening software products still require substantial manual oversight from highly trained experts to monitor and sort hard-to-read sequence hits, particularly false positives. The associated labor costs and the computational intensity of current products make screening expensive and hard to scale.

In a worst-case scenario, high screening costs would give companies who skimp on safety a competitive advantage over those who implement more stringent safety best practices. Indeed, this is what industry aims to avoid. In 2023, the **CSIS Bipartisan Alliance for Global Health Security**—a biosecurity consortium including senior members of Congress, industry experts, and biosecurity scholars—**noted** that "private industry is reportedly asking for mandated screening of DNA synthesis orders to ensure responsible actors are not financially penalized for implementing safeguards of their own volition." (See Appendix C for upcoming enforcement provisions under the 2025 Trump executive order.)

More recently, the April 2025 **NSCEB report** states that "while industry is united in calling for a measured, enforceable, and standardized approach [to synthesis screening], the U.S. government is unable to respond to such requests with the needed agility" to make this happen. This is because current government policies are split between multiple federal agencies, "leading to redundancies, gaps, and inefficiencies" in biosecurity, including U.S. government screening guidelines.[13] Thus, the commission concludes that current U.S. biosecurity measures place "undue burden on researchers and innovators to navigate unwieldly bureaucratic processes while enduring market and academic pressures."

Unfortunately, biosecurity measures such as synthesis screening are fragmented at a global level as well. For now, U.S. companies **dominate** the global synthesis market, yet international providers occupy a growing share. U.S. customers could thus exploit the **lack** of international synthesis security standards and choose foreign providers that employ more relaxed measures to produce dangerous sequences abroad. Foreign actors could similarly choose firms in countries with fewer government guidelines, if any, to synthesize their sequences. Standardized synthesis requirements for U.S. companies alone are therefore not a catchall solution to addressing international biosecurity challenges, especially given that epidemics and pandemics can naturally transcend borders.

Appendix E

Additional Considerations for an AI-Enabled Synthesis-Screening Process

BUILDING ON EARLIER U.S. GOVERNMENT RESEARCH: IARPA'S FUN GCAT PROGRAM

In looking to develop an effective synthesis-screening process, government and industry stakeholders need not start from scratch. Previous government-sponsored research demonstrates that an AI-enabled screening tool is not only possible but has been achieved before.

From 2017 to 2022, the Intelligence Advanced Research Projects Activity (IARPA) ran a program to develop an AI-enabled DNA sequence-screening software that would be able to predict sequence functions and gauge potential threats. This **Functional Genomic and Computational Assessment of Threats** program aimed to replace what its developers called "inadequate" screening methods with a suite of computational tools that could assess dynamic and emerging biological threats. As one 2016 IARPA document introducing the program rightly **stated**, "We have little or no capacity to deal with novel, emerging, or unknown sequences." Thus, what the United States needs is to "be creative, novel and forward looking" by taking "experimental approaches to enable understanding risks of genes" using "advances in computational prediction of [genomic] structure and functions."

According to IARPA, the program demonstrated significant success while it was operational. As the program website states:

> Fun GCAT tools are demonstrating high predictive accuracy against increasingly challenging test sets. Benchmarking has demonstrated significant performance increases beyond top winners in a closely related bioinformatic software development global challenge. . . . This enabled a range of Intelligence Community-relevant missions from rapid screening of very large datasets to field-based, targeted analysis.

IARPA's Fun GCAT program concluded in 2022, but experts told CSIS that it remains well-recognized in the biosecurity community today as a successful program for improving nucleic acid synthesis screening. It offers a proof of concept for an AI-enabled screening model, as well as evidence that the U.S. government already thinks such an idea is worthwhile. The model proposed in this paper would build upon IARPA's project with the aim of making it available to commercial synthesis companies.

SECURITY CONCERNS AND POSSIBLE MITIGATION MEASURES

For an AI model to be able to screen dangerous sequences effectively, it would need to have the necessary harmful information in its training dataset. This would entail leaving out many of the biosecurity-related guardrails common to commercially available models.

Because a model of this kind would have dangerous dual-use capabilities, careful consideration is needed regarding model access and ownership. DNA synthesis providers should be able to screen every order using this AI model and do so quickly and cheaply. Providers should thus be granted easy access to the model. However, because of dual-use concerns, only certain credentialed actors should be able to possess an actual copy of it.

Government and industry stakeholders would thus need to develop a mechanism that reconciles the ease of access for providers with strict security controls. One solution could be to store the model in a federal agency-secured cloud—perhaps hosted by the National Institutes of Health, the Department of Defense, or the Department of Commerce—that synthesis providers ping via encrypted channels to screen sequence orders. The model could scan the sequence then automatically send a risk assessment back to the provider (see Figure 7). Allowing only authorized U.S. government officials access this AI model could reduce the likelihood that the model is misused, tampered with, or overridden by malicious actors.

OUTSTANDING CHALLENGES

Though AI biology models present an exciting opportunity to strengthen U.S. biosecurity, there are limitations to this proposed AI-enabled screening measure. For one, it would only apply to commercial DNA synthesis firms. While these firms currently dominate global DNA synthesis production, a growing number of smaller "benchtop" DNA synthesizers could make it easier to circumvent screening entirely by allowing sufficiently trained users to produce sequences themselves. Though the technology is still maturing, the NSCEB estimated in 2024 that benchtop devices will be able to synthesize DNA sequences the length of the smallest viruses in the next two to five years. As a December 2024 report from the Institute for Progress highlights, expected rapid advancements in benchtop synthesis will pose U.S. biosecurity risks that current government measures are ill-equipped to manage.

Another outstanding limitation that does not yet have a clear solution is the risk of malicious actors ordering fragments of sequences across multiple companies with the intention of recombining them later. Unfortunately, there is no obvious fix for this issue that would not require excessive compliance burdens for legitimate and benevolent actors. However, the AI Action Plan recommends that OSTP lead the development of a mechanism for synthesis providers to share data regarding customer screening, a recommendation that, if effectively executed, could help to address some of this issue.

Finally, the success of this proposed measure clearly requires the continued advancement of biological AI capabilities beyond the current state of BDTs and other AI models. It also requires funding for relevant federal agencies and a clear delineation of agency responsibilities relating to AI and biosecurity. For example, as the 2025 NSCEB report discusses in depth, fragmented agency authorities relating to biosecurity is especially challenging for developing and implementing effective policies. The NSCEB advocates for a centralized agency to oversee U.S. biosecurity policy, a recommendation that the authors of this paper support.

Unfortunately, deeper analysis of these complex issues is beyond the scope of this initial paper, and it is why the authors recommend developing only a plan to further integrate AI into synthesis screening. The efficacy of this proposal necessitates close partnerships between U.S. policymakers, industry, and academia on numerous considerations, including but not limited to technological feasibility, industry incentives, and security control measures.

Therefore, an AI-enabled synthesis-screening process should serve as just one component of the wider U.S. biosecurity toolkit rather than a comprehensive solution.

Endnotes

[1] This paper defines "bioterrorism" as the deliberate release of biological weapons to cause death or disease among civilians by nonstate actors. Other related biological risks, such as state-sponsored biowarfare and accidental misuse, are part of the wider biological threat landscape of which bioterrorism is one component. Though some biosecurity strategies discussed in this paper may help address these related risks, the following analysis focuses on disrupting bioterrorism capabilities specifically. The authors acknowledge that different stakeholders use the word "biosecurity" to cover varied spectrums of risk.

[2] Biological design tools, or BDTs, are broadly **defined** as AI models trained on biological data such as genetic sequences specifically for biological research and development applications. This report refers to AI biology models such as Evo 2 as BDTs to avoid confusion with popular large language AI models discussed elsewhere. It uses the terms "model" and "tool" interchangeably when referring to BDTs to reflect the common terminology of their developers.

[3] The precise number of technical experts working in Aum Shinrikyo's biological weapons program is unknown, but approximately **300 scientists** are believed to have worked on the group's biological, chemical, and nuclear weapons programs combined. CSIS conversations with experts familiar with the subject indicate that Aum Shinrikyo's bioweapons program may have included a dozen technical experts or fewer.

[4] Technological limitations in DNA synthesis require synthesis companies to produce fragments of genetic sequences instead of entire genomes. This is because manufacturing beyond the length of **5,000 base pairs** (the fundamental unit of DNA) introduces drastically higher sequence errors, complexities, and costs. The horsepox genome has approximately **212,000 base pairs**, meaning that the Canadian scientists would likely have had to break up their order into at least 43 separate fragments.

[5] Note that many experts believe that the mortality rate of bird flu (and especially the H5N1 avian influenza) in humans is likely overstated because mild cases often go **undetected**.

[6] For instance, the NASEM report **acknowledged** that "the de novo design of a virus would represent a significant capability uplift enabled by AI biological tools . . . [and] could have the highest impacts in terms of consequences," but does not expand upon this assessment.

[7] Note that while AI can make developing new harmful sequences more straightforward, this capability has been within reach of more sophisticated actors for years. As early as 2006, for example, the **National Science Advisory Board for Biosecurity** warned, "It is now feasible to produce synthetic genomes that encode novel and taxonomically unclassified agents with properties equivalent to, or potentially more harmful than, current Select Agents."

[8] These frameworks complement broader U.S. government efforts to strengthen biosecurity, such as the Biden administration's 2024 **Policy for Oversight of Dual Use Research of Concern and Pathogens with Enhanced Pandemic Potential**, the Department of Defense's 2023 **Biodefense Posture Review**, and the National Security Council's 2022 **National Biodefense Strategy and Implementation Plan**.

[9] Depending on how one accounts for different taxonomy/classification/nomenclature conventions, the precise number of unique items on the BIS Commerce Control List could range from roughly 130 to 180 items. CCL's dual-use export control objectives mean that it demonstrates much broader regulatory coverage than the HHS/USDA Select Agent and Toxins List, which is focused more narrowly on a small selection of the most severe biological threats to human, animal, and plant health in the United States.

[10] For a more detailed discussion of which BDTs should be tested by the U.S. government and how to define these thresholds, readers are encouraged to consult Doni Bloomfield et al., "**AI and Biosecurity: The Need for Governance**," *Science* 385, no. 6711 (August 22, 2024): 831–833, doi:10.1126/science.adq1977.

[11] Since the 2024 OSTP framework has been removed from the U.S. Administration for Strategic Preparedness and Response (ASPR) **website**, it is unlikely that this target date and other requirements remain in effect following President Trump's **revocation** of the AI executive order on January 23, 2025, and subsequent issuance of a new **executive order** mandating the revision or replacement of the 2024 OSTP framework within 90 days (see Appendix C).

[12] Experts **disagree** on the **feasibility** and even desirability of developing a common SOC database. Some argue, for instance, that such a database could itself pose biosecurity risks by centralizing all known information about SOCs, which could in turn simplify malicious actors' search for harmful genetic information should the database's security controls be breached. Others argue that while a centralized U.S. government-developed database could help standardize screening processes, the government lacks the authority and funding to make this happen. These debates are ongoing among industry and experts, indicating that a definitive SOC database is unlikely to arise anytime soon.

[13] A company adhering to the 2023 screening framework, for instance, would likely need to consult with at least five separate government agencies, including HHS, the International Trade Administration, the Federal Bureau of Investigation, the U.S. Department of Commerce, and the National Institutes of Health.