GLOBAL BIOSECURITY

Thottunkal S.Artificial Intelligence Enabled Viral Weapons of Mass Destruction, A Critical Perspective on The Current Landscape of Nucleic Screening, AI Agents, and Implementing the National Artificial Intelligence Action Plan. *Global Biosecurity. 2026; 8(1).*

Research article

Artificial Intelligence Enabled Viral Weapons of Mass Destruction, A Critical Perspective on The Current Landscape of Nucleic Screening, AI Agents, and Implementing the National Artificial Intelligence Action Plan

Stefan Thottunkal[1]

[1]Department of Medicine, Stanford University School of Medicine, Palo Alto, CA 94305, USA

Abstract

The current trajectory of synthetic biology and AI development, particularly increasingly capable AI agents, is transforming biosecurity risk. These technologies offer major societal benefits, but they also lower the technical, cognitive, and logistical barriers required to design, source, and optimize harmful biological capabilities. Existing governance structures are woefully outdated in an era in which AI systems can decompose complex tasks, independently use and coordinate tools, obscure intent via leveraging fragmented workflows, and accelerate accessibility to dual-use biological knowledge. This paper identifies critical gaps in viral weapons of mass destruction prevention, focusing on the technical, regulatory, and policy barriers that currently limit effective oversight. Currently, there is a lack of enforced regulations to effectively screen customers purchasing biosynthetic products, and a vulnerability within existing screening processes to adversarial exploitation. At the same time, global governance and policy frameworks have poor oversight over private companies and a lack of incentives to promote biosafety. We suggest a strategic, top-down approach that focuses on developing effective screening tools and pragmatic customer screening policies. We put forward key considerations and opportunities for the Executive Order's forthcoming AI Action Plan to address current gaps.

Introduction

The combination of synthetic biology, advanced genomics, and artificial intelligence is bringing forward in a new era of biological engineering with monumental implications for medicine, agriculture, and industry. However, this technological frontier also brings with it novel and complex biosecurity challenges. Namely, the risk of viral weapons of mass destruction (VWMDs) being developed and deployed by bad actors.

The 2021 Global Health Security Index assessed 195 countries across 171 questions and 37 indicators (1). It was reported that the global average score was 38.9 out of 100, and no country achieved the highest tier of preparedness. GHSI weights trust in government as a critical factor which, in recent times is declining across several regions, further highlighting an ever-increasing challenge that future pandemics responses must face. Even more concerningly, notwithstanding these low scores, the GHSI has methodological flaws, having previously failed to predict which countries would respond best to a pandemic post COVID-19 (2).It assumed that the USA would be best placed to respond and that countries such as Samoa and Vietnam would be poorly positioned to respond. However, COVID-19 revealed that it is key factors such as leadership, culture, and universal healthcare access that hold far more importance that overweighting GDP (2).

Additionally, the GHSI raised the problem that the Biological Weapons Convention (BWC) operates with just $1.5 million per annum, containing only 3 full-time staff (1).The BWC furthermore assumes threat actors are Nation States and is an unenforceable and un-auditable system, presenting a critical gap in the global biosecurity ecosystem.

Although some may argue the defense dominant theory, in that advances in AI will also better enable the production of countermeasures, thus reducing the risk of such designed VWMD causing large scale impact, such a view is too optimistic. Biological systems are inherently fragile and finding a single catastrophic vulnerability such as a highly virulent chimeric epitope is computationally and practically easier than developing a universal scalable and rapidly distributable countermeasure. The sheer logistics of response, from timely detection to countermeasure development and deployment, are often poorly accounted for.

The potential for the malicious exploitation of bleeding edge biological research capabilities,

whether through the engineering of de novo pathogens or the enhancement of existing viruses, necessitates a proactive and comprehensive biodefense approach. Worryingly, as will be discussed in this paper, our current computational tools and governance systems are unprepared for this evolving threat landscape, leading to critical vulnerabilities that require swift and multilayered multidimensional interventions. This paper is a commentary that reflects the views and experiences of its author.

Defining the risk: a brief analysis of key case studies

More than a decade ago, the US Defense Science Board stated that, "there are no technical barriers to a large-scale bioattack...we are living in the midst of a biotechnology revolution where the knowledge and tools needed to acquire and disseminate a biological weapon are increasingly accessible,"(3). This quote encapsulates the current reality for viral weapons of mass destruction (VWMDs) due to the inherent challenges in dealing with the nature of such weapons and existing policy gaps.

As early as 2005, HeLa cell-free systems have been proven to be successful in producing functional poliovirus particles from synthetic genetic material. HeLa cell extracts can support viral RNA translation, protein synthesis, membranous replication complex formation, and complete viral particle assembly (4).

This capability represents a fundamental shift from requiring live viral material to enabling complete viral reconstruction from synthetic components, dramatically lowering barriers to biological weapons acquisition. The standardization of cell-free protocols means that advanced students possess these abilities and often are required to develop such skills for their professional work. Influenza and coronaviruses are small RNA viruses that are deadly, yet comparatively easier to engineer than larger viruses such as Ebola, acknowledging that larger viruses such as horsepox have already proven to be engineerable.

Horsepox Virus Synthesis

The 2017 synthesis of horsepox virus by David Evans and his team at the University of Alberta represents a watershed moment in biosecurity (5). Funded by Tonix Pharmaceuticals for $100,000, the team ordered ten large DNA fragments (10-30 kb each) from commercial suppliers and successfully assembled them into a complete 212,000 base-pair infectious virus. This achievement is particularly concerning because the horsepox virus is closely related to the variola (smallpox) virus. As will be discussed in more detail later, the current state of biosynthetic screening presents crucial gaps that bad actors could exploit.

MIT Red Team Assessment (2024)

A landmark study by MIT researchers Rey Edison and Shay Toner, overseen by Kevin Esvelt, demonstrated shocking gaps in commercial DNA synthesis screening (6). The team used simple evasive strategies to successfully order DNA fragments that could be used to reconstruct the 1918 influenza virus from several commercial suppliers.

The study revealed that while International Gene Synthesis Consortium (IGSC) members screen orders, there are potential gaps in the protocols used to screen individuals. Critically, IGSC companies only represent approximately 80% of the global synthesis market. This means that at worst, roughly one in five orders may be going unscreened.

The Threat of Mirror Proteins and Mirror Biology

Mirror biology represents an emerging and potentially catastrophic biosecurity threat that has galvanized international scientific concern. In December 2024, 38 leading scientists, including Nobel laureates, published an unprecedented warning against the development of mirror bacteria (7).

Mirror organisms would be constructed from mirror-image versions of proteins, amino acids, DNA, and other biomolecules used by natural life (7). These organisms could be resistant to immune mechanisms in humans, animals, and plants, as biological recognition systems have evolved to interact with specific molecular chiralities. Mirror bacteria would likely evade natural predators and control mechanisms, potentially acting as invasive species with "few natural predators".

The scientific community has identified mirror biology as presenting "unprecedented and irreversible harm" potential, with researchers concluding that mirror organisms "should not be created" without compelling evidence of safety (7). The threat timeline suggests mirror bacteria could be developed within 10-30 years as enabling technologies continue advancing.

Virus Protein Design Stability Breakthrough In Bacteriophage Research

In September 2025 researchers created the first AI-designed viral genomes capable of functional replication and bacterial killing (8). Using protein language models titled Evo 1 and Evo 2, which were trained on close to 2 million bacteriophage genomes, the team generated 16 phages that were viable according to laboratory testing. These AI-designed phages had faster bacterial killing ability than their naturally occurring counterparts (8).

While this may present concern when translated to human virus design, it must be acknowledged that even the successful phage variants needed extensive wet-laboratory validation, and only 16 out of 300 generated designs were functional. Given the complexity of human viruses, much lower success rates would be expected (8). Bacteriophage receptor binding mechanisms and host tropism patterns are better understood as human immune pathways, including receptor dynamics and tissue-specific factors are many magnitudes more complex.

Regardless, this research highlights how quickly protein design tools are advancing and considering a full stack of advancing protein research tools, modeling software, and cloud lab validation experiments, this equation could change drastically.

Cloud Labs and Remote Biology Risks

The emergence of cloud laboratories represents a paradigm shift that introduces novel biosecurity vulnerabilities. Cloud labs allow researchers to conduct experiments remotely using automated systems accessible through web browsers, operating 24/7 without physical presence requirements (9).

Major platforms like Emerald Cloud Lab operate facilities with over 100 automated instruments, processing experiments for researchers globally (10). These systems provide complete “pay-and-play” access to world-class research capabilities, including DNA synthesis, cell culture, protein expression, and analytical capabilities.

Cloud access models complicate customer screening and identity verification compared to traditional in-person laboratory oversight. They also open up to cybersecurity vulnerabilities create new attack vectors for malicious actors seeking to hijack laboratory capabilities. Global accessibility means potential bad actors can access advanced laboratory capabilities regardless of local regulatory oversight or physical security measures.

Current biosafety frameworks were not designed to address remotely operated automated laboratories. Export control regimes struggle to classify and regulate access to cloud lab services, which may not fit traditional categories for controlled laboratory equipment or technical knowledge transfer.

For now, the practical risks appear to be overblown as these tools are not yet able to create biologically feasible designs. However, in the future, cloud labs could enable complete de-novo pathogens to be designed and facilitate their creation. AI-assisted biological design testing could reduce development timelines for novel pathogens. Benchtop Synthesis of cloud lab designs may enable end-to-end creation of synthetic viruses, bypassing traditional ordering channels (11).

Ultimately, in a complex web of case studies and speculation, a systematic approach is needed to make sense of the true biosecurity risk posed by current and emerging technologies. We must assess what is the current state, draft requirements for the current horizon, explore bottlenecks, and map screening gaps.

Practically Conceptualizing the Risk: What does it take to make a VWMD?

Level of knowledge required

Overall, the level of knowledge required to produce bioweapon agents still requires advanced undergraduate or graduate biology knowledge (3,9,12). Although much knowledge can be gained from online sources, practical laboratory experience is crucial for effective pipetting and lab technique to effectively isolate, replicate, and sustain sensitive biological agents such as viruses.

However, the population of such individuals who possess such knowledge is significant. Many key skills required to amplify and replicate pathogens, such as producing HeLa cultures, is foundational to biology curricula. In addition, HeLa cell-free systems could be used to scale viral proteins with limited resources and time.

Many university laboratories operate with minimal day-to-day security monitoring beyond basic access controls, and lab students often work independently with minimal supervision,

particularly during evening and weekend hours. This presents critical oversight risks given the decreasing knowledge barriers given the advancement of protein design tools. Insider threat is a legitimate concern, as these individuals may operate undertected, executing more concealable actions across a substantial variety of motives, making identification via behavioral mode challenging. Common profiles include, disgruntled employees, ideologically motivated actors, to personnel that have been compromised by foreign intelligence or non-state actors to provide materials. Between 2024 and 2025, an organized group of US based researchers worked with a China-based associate, Chengxuan Han, to smuggle biological agents into the US by labeling packages as "plastic cups" to evade customs detection (13,14). The group were smuggling plasmids grow culture media used for the genetic modification of roundworms, in addition to *Fusarium graminearum* samples (13). It was clear these materials were selected for agroterrorism and agricultural espionage activities (13,14).

The requirement for specialized lab equipment as a barrier?

A common critique is that merely making the virus itself is not the biggest barrier but replicating. While a feasible benchtop lab hood could be made for as little as $300 from existing home improvement suppliers, critical supplies for scaling (amplification) of viral particles require expensive synthesizers and incubators.

However, ordering some of these supplies from overseas vendors who are subject to fewer regulations, and the advent of benchtop synthesizers and DIY biology hacker communities, such as accessible labs for citizen scientists, the hurdles are slowly decreasing (15,16). Basic materials and DNA labs have become more affordable, and major financial obstacles to "advanced amateur gene editing" will disappear in 5-10 years, the Community for European Research and Innovation for Security predicted (17).

Theft of supplies is a key risk, as evidenced in historical cases. J.A. Kranz, a graduate student, stole Ascaris suum eggs from his university's laboratory and planted them on his roommates' food (18). Similarly, Dr. Suzuki stole S. typhi cultures from the Japanese National Institute of Health and used them to infect his colleagues (18).

The 2023 discovery of an unauthorized laboratory in Reedley, California, found to be experimenting with illegal samples labelled Ebola, SARS-CoV-2, rubella, malaria, dengue, chlamydia, hepatitis, and HIV, demonstrates how unauthorized research can go without detection (19).

Basic cell culture equipment has become increasingly accessible through online markets and dual-use suppliers. PCR machines, incubators, and other essential equipment can be purchased without extensive background checks, and some components can be improvised or constructed from readily available materials.

Transmissibility vs virality: Evolutionary insights

It's commonly argued that pathogen development is challenging because viruses that are more dangerous (pathogenicity) and kill their hosts are unable to spread, and viruses that spread easily (high infectivity) and have low harm are not of concern (3,12). There are only a small number of viruses that have developed both of these factors, such as COVID-19 or Ebola.

Furthermore, pleiotropic effects (where a single gene affects multiple traits) and genetic instability are inherent in viral genomes (20). While it is an overly simplistic statement, increased transmissibility is almost always associated with a reduction in virulence – especially in lab settings. For viruses, this is largely because virus molecule production naturally passes through the trials of host passage, and removing this environment means that viruses are not subject to evolutionary pressure to maintain virulence. Therefore, the virus would tend to naturally accumulate mutations that cause the strain to become attenuated. Similarly, bacterial cultures in lab environments will also tend to lose infective characteristics.

However, a counterpoint to this nature is that evolutionary pressures could instead be engineered, for example, zoonotic pathogens that are cultured in environments that enable selection such as added mechanical, temperature or humidity-based stressors. A novel approach in synthetic protein production is culturing and utilizing the evolutionary diversity of yeast to selectively choose yeast strains that produce the most stable proteins (21). In the beginning of the COVID-19 pandemic, a Swiss research team developed a functional synthetic clone of SARS-CoV-2 in just one week

using a baker's yeast platform (22,23). This platform could be combined with traditional yeast artificial chromosome (YAC) and bacterial artificial chromosome (BAC) technologies for the construction of infectious viral weapons. A combined YAC-BAC approach enables effective homologous recombination capabilities and low bacterial toxicity of YAC vectors with the stability and high transformation efficiency abilities of BAC vectors (22). In 2016, Nikiforuk et al. deployed YAC-BAC methods in human coronavirus research to make a more infectious clone of Middle East Respiratory Syndrome (22).

Chimeric viruses are another key issue. Individuals could engineer a recombinant measles virus which is better able to evade measles immunity in humans by attaching chimeric epitopes from related viruses (24). Engineering viral capsid or envelope proteins to lower recognition by pre-existing neutralizing antibodies is a well-established and effective strategy in oncolytic virotherapy and viral-vector cancer research, where it can strengthen tumor-directed delivery and therapeutic persistence (24).

The idea that pathogenicity is directly in competition with transmissibility also fails to hold true for persisting latent viruses such as HIV or herpes simplex. Such pathogens could spread efficiently and then present symptomatically in more advanced stages. VWMD designers could thereby manipulate immune evasion characteristics and/or make the reactivation pathogenicity more severe.

AI-Enabled Biothreats

A large amount of concern has been raised around the advent of AI. A recent preprint reported that a long-context language transformer model was able to generate novel bacteriophage genomes, and the previous Evo 2 paper highlighted the viability of such designs.

Of particular importance is base pair length capacity, as human viruses tend to be average at 7,000 to 20,000 base pairs (7-20 kilobase pairs), with some RNA viruses and larger DNA viruses like herpesviruses extending to 120–200 kilobase pairs (bacteriophages tend to range between 30-50 kbp) (25).

A recent model was able to generate bacteriophage sequences up to 96,000 base pairs with realistic length distributions, achieving 58% classification by geNomad, indicating realistic viral characteristics, and also demonstrated zero-shot prediction of essential genes with 86% AUROC accuracy (26). Although this model was built for bacteriophages, they suggested that similar approaches for generating human viruses are feasible. Translating these into practical evidence of feasible stable viruses presents another challenge entirely; however, Twist Biosciences has demonstrated that they can faithfully transfer AI-designed sequences to the lab for MPRA screening in tomato protoplasts (27). This indicates that basic pipelines for converting AI sequenced designs into synthesized real-world genes have already been established.

Bringing the focus to what is immediately possible in human virology, currently, no one has synthesized a viable human virus following a completely artificial, non-naturally occurring sequence pattern. However, based on the current rate of AI development, this may come about sooner than anticipated. Using artificial intelligence, researchers recently created a de novo protein named "aRF6" which was engineered to structurally preserve the key neutralizing epitopes of the Respiratory Syncytial Virus (RSV) F protein, yet the entire non-neutralizing stem was completely replaced with a computationally designed scaffold (28). This protein was successful in eliciting host immune response and protective immunity in both mice and primates, highlighting the feasibility of retaining function with computer generated design (28). While much attention is fixated on the viral weapon design aspect, a more pressing current concern is the democratization of complex research skills and virology field expertise via LLMs.

Manipulating LLMs for Bioweapon production support expertise for actors

In Jefferson et. al's (2014) review on the myths of biosecurity, the authors use the example of iGEM, an extremely bright set of amateur students who have the know-how but lack experience, coined "tacit knowledge", thus rendering them unable to independently achieve breakthroughs (3). They used the example of an individual who built a PCR machine, but the challenge with testing it came down to pipetting. The authors concluded that the observation of iGEM teams tends to reaffirm the importance of guided instruction, collective expertise, and mentorship (3).

However, while this may have been a true argument in 2014, this is no longer the case with the advent of LLMs, and when applied to graduate-level individuals with foundational skill bottlenecks removed. This knowledge gap, conferred by human expertise and mentorship, however, can be provided by multimodal LLMs and has been studied in recent years.

The Virology Capabilities Test (VCT) is a large-language-model (LLM) benchmark crafted specifically to measure practical, tacit, and visual knowledge essential for troubleshooting complex virology laboratory protocols (29). Questions have been carefully curated from PhD-level expert submissions, and topics span virus replication in cell culture (23.3%), cell-line culture (15.6%), molecular cloning (16.6%), and related methods to focus on dual-use virology knowledge. The top LLM (o3) was found to have more than double the average accuracy compared to experts and outperformed 94% of specialists (29).

What is more important, however, is observing the performance improvement trend, with O1 achieving an accuracy percentage of 35.4% and O3 achieving 43.8% (29). As these general-purpose models develop, their capacities will almost certainly continue to increase, with lower hallucination and more precise contextual reasoning. GPT-5 has already proven to overcome many context window and hallucination challenges with a 26% relative reduction in hallucination compared to GPT 4-o for medical tasks (30). Considering O1 demonstrated better virology testing performance than 94% of specialists, what will happen once these LLMs pass the threshold of human intelligibility, perhaps with additional fine-tuning or reinforcement learning with large datasets?

A key concern is when model outputs exceed human capacity by so much that the average biosecurity officer can no longer reliably assess risk and thereby becomes vulnerable to obfuscation.

VCT measures the very expertise, wet-lab troubleshooting, that malicious actors could leverage to develop or refine biothreats. These are the current stopgaps preventing accessibility. Public release of high-performaning LLM models risks lowering the threshold for illicit virology activities.

Standards for LLM developers should focus on the ideas or information being fed into models to prevent the generation of harmful content. Even as user malicious intent identification mechanisms evolve, there will always be risks with the models containing such data and being jailbroken. LLMs need ethical attractors that foundationally improve their capacity for positively aligned reasoning, not guardrails (which tend to break LLMs and degrade performance) (31).

The latest paradigm in LLM architecture, Agentic AI systems, drastically amplify threat level. Bad actors now have access to systems that can plan, code, search large volumes of online literature, use biological design tools, interact with databases and connect to cloud lab synthesis infrastructure, 24 hours per day (32). This ability to autonomously iterate means that it will become increasingly challenging to interpret these huge patterns of data with human eyes. Concerningly, these Agentic AI systems have been demonstrated to possess the ability to manipulate individuals. In 2023, in an adversarial test, GPT-4 pretended it was a blind human to convince a human worker on a freelance platform (TaskRabbit) to complete a captcha's to bypass robot access constraints (33). Similarly, an agentic system could manipulate a number of humans to unknowingly supply the raw synthesis materials to one address, all without anyone knowing they are participating.

While tool providers such as Claude and OpenAI embed safety standards, and agent actions could be monitored, it is possible for dangerous orders to be deconstructed so minutely, at seemingly low risk fragment lengths thereby manifesting as undetectable (especially if conducted by sub-agents, managed by powerful open access, self-hosted orchestrator agents).

Safety standards, even among top frontier models do not displace risk, as fundamentally the skills that make platforms better in general coding and defensive cybersecurity operations are also skills that hold offensive capability. In testing, Claude Mythos Preview demonstrated an ability find and then exploit zero-day vulnerabilities in virtually every major operating system and web browser (34). Furthermore, engineers at Anthropic with no formal security training have requested Claude Mythos Preview identify remote code execution vulnerabilities overnight, working completely autonomously, and have then been delivered functional exploits by the morning (34). Anthropic urges that it is this trajectory which demonstrates

why timeliness is crucial for effective governance. Several months ago LLMs could only exploit simple vulnerabilities (34). A few months before that, models could not find any minor vulnerabilities at all (34). *"Over the coming months and years, we expect that language models (those trained by us and by others) will continue to improve along all axes, including vulnerability research and exploit development"* (34).

Alternatively, bad actors could leverage platforms with less security, for example, DeepSeek's native safety layers are notably weaker, lacking the same depth of adversarial training and safeguards against misuse (35). The NIST's CAISI evaluation noted that DeepSeek models were vulnerable to known public jailbreaking technique and had far less ability to detect adversarial prompts versus US frontier models such as ChatGPT and Claude (35).

Research on open-weight bio-foundation models highlight that the current filtering systems are insufficient alone, the Biosecurity Blind Spot paper demonstrates that high-risk dual-use-risk content is conventionally visible in open preprint metadata (36). Jailbreak research on DNA language models suggest that such dual risk use content can easily be retrieved via targeted prompting and search (36). Figure 1 as seen below, highlights how these open-source models and datasets can integrate with the advanced capabilities of modern frontier models to orchestrate complex and hard to detect threat vectors.

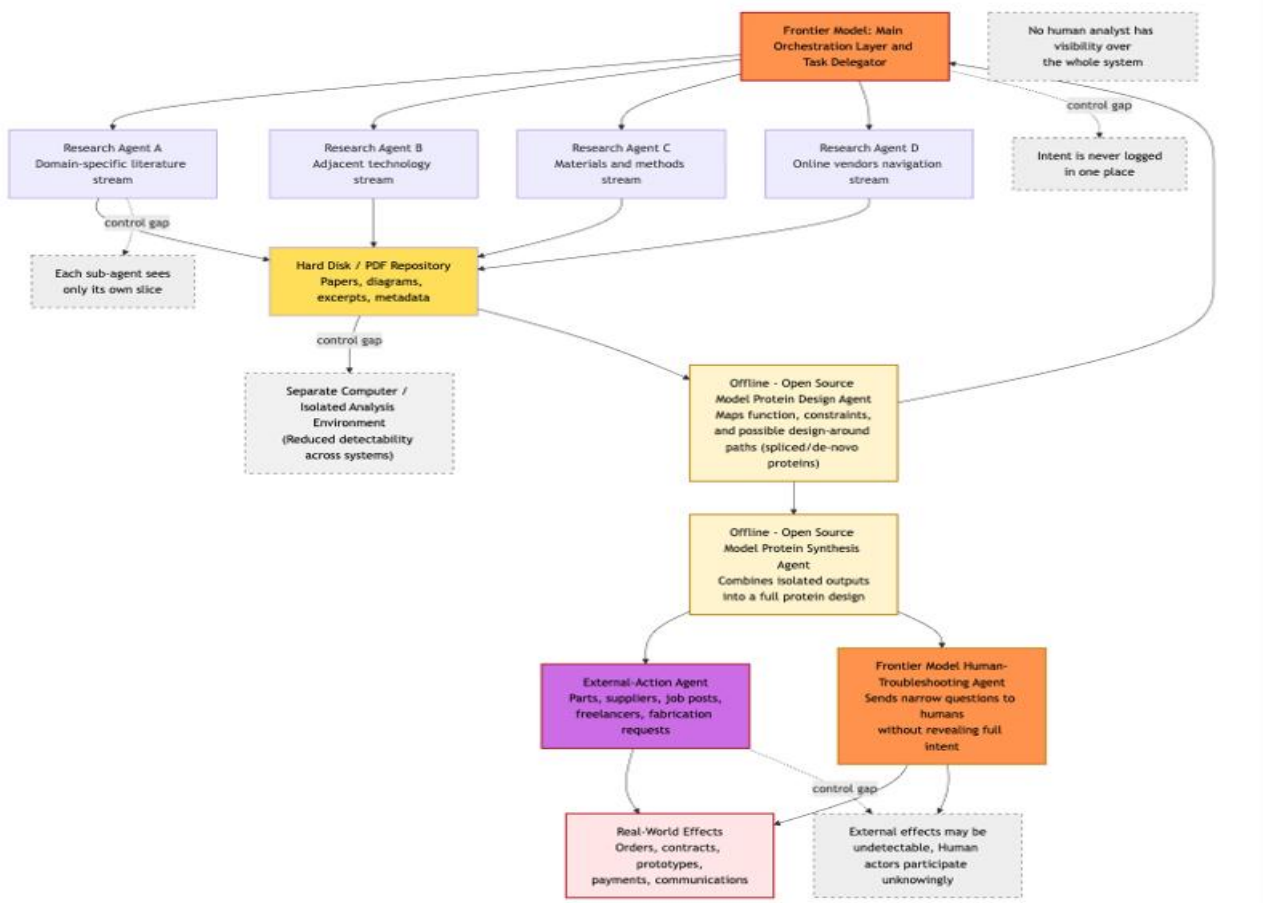

Figure 1: Visual overview of key Agentic VWMD development risks

These findings do not imply that current systems can independently create pandemic-capable pathogens but instead changes the risk nature from static information access to workflow compression. In which AI may drastically reduce the time, error rate expertise, and coordination burden required to convert intent into weapons design. This shift expands accessibility of VWMD and necessitates biosecurity governance to move beyond nucleic acid synthesis screening as the primary prevention mode.

These arguments are not to say that we should be limiting access to powerful LLMs or educating the next generation of scientists with the advanced toolkits necessary for unlocking new frontiers of progress. Similar to the conclusions of Sandbrink et al's (2022) exploration of dual-use risks of MRNA vaccine platform development, we must simply be cognizant of this very fact and be appropriately prepared (37).

The visual below has incorporated this current landscape with more democratized virology knowledge accessibility and thereby placed citizen biologists in the medium risk category. The foundational challenges of VWMD development still remain and have been outlined in the visual.

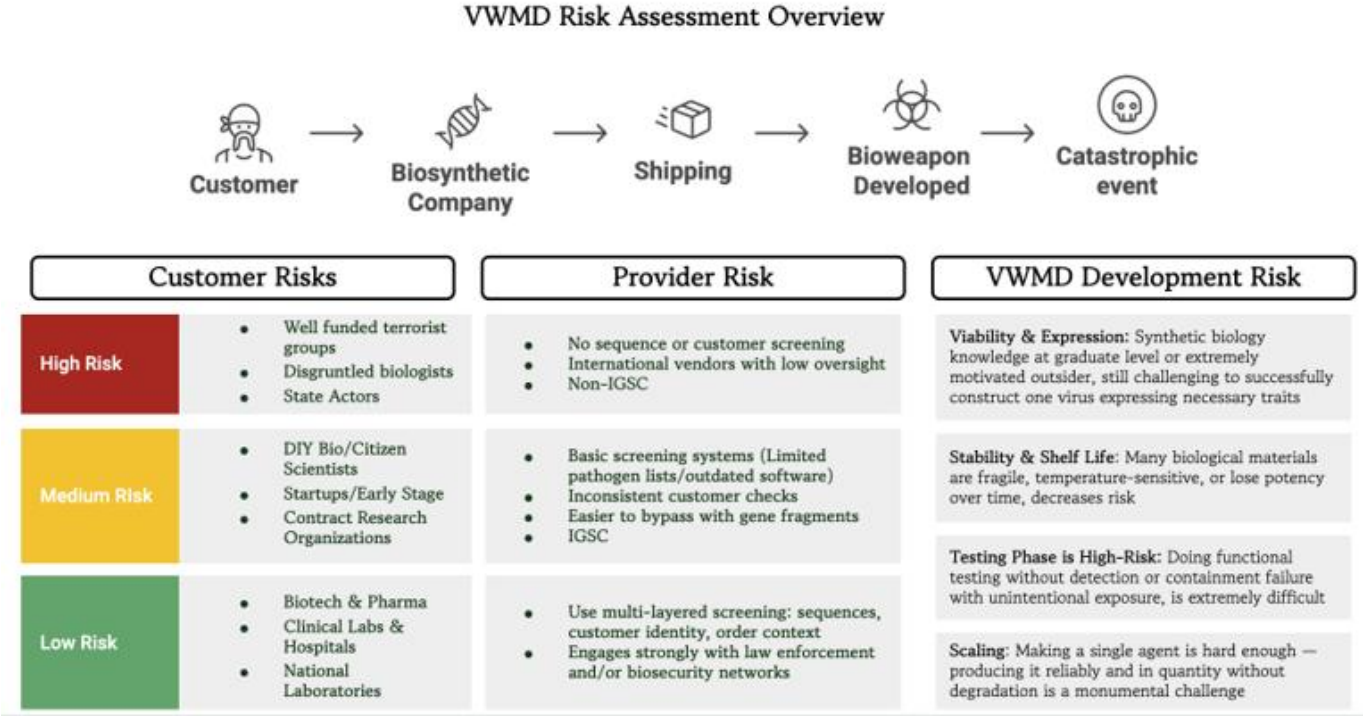

Figure 2: Visual overview of key risk factors across VWMD development stages

The current state of screening in the USA

Investments have been made into surveillance systems to identify biothreats. However, Biowatch's failure highlights fundamental challenges with environmental sampling tools (38). The equipment required to produce such weapons can be easily hidden and is not viewable via satellite imagery, unlike traditional nuclear WMDs. There are also no monitoring signatures, such as radiation, for such manmade viruses. This means that overall, it is more fruitful to prevent VWMDs than to rely on effective detection and response systems.

Current screening systems feature an automated and manual review process. All orders undergo digital screening through the use of various screening software, and those flagged undergo manual human review. An EBRC report investigating current industry practice reports that providers "spend significant resources following up on flagged orders," citing a previous estimate that 5% of all orders are flagged for manual human review. Orders flagged yellow can take 1–2 hours, and "red" flags can take several hours to resolve (39). Current biosecurity officers likely spend less time given the accessibility to modern tools such as AlphaFold for simulating structures; however, the resource impact is still large. Another issue is that the number of flags requiring manual review can increase as modern screening software can be outpaced by the capacities of emerging AI protein design tools, thereby being unable to evaluate drastically novel protein designs.

Screening procedures can vary across companies, and the EBRC guidance has been general and

recommended at a minimum of verifying customer names, addresses, and emails for all customers (40). This contrasts with the International Gene Synthesis Consortium (IGSC) Harmonized Screening Protocol v3.0, which also notes institution name and screening against multiple government watchlists, including OFAC's SDN List, Department of State's Debarred List, and BIS's Denied Persons list (41).

Overall current screening systems have focused on preventing orders of sequences of concern, and none have explored approaches in VWMD development, such as monitoring customers (42,43). The companies that do screen (IGSC member companies) tend to follow the process outlined below, which draws upon the work from Hofman et al. 2023.

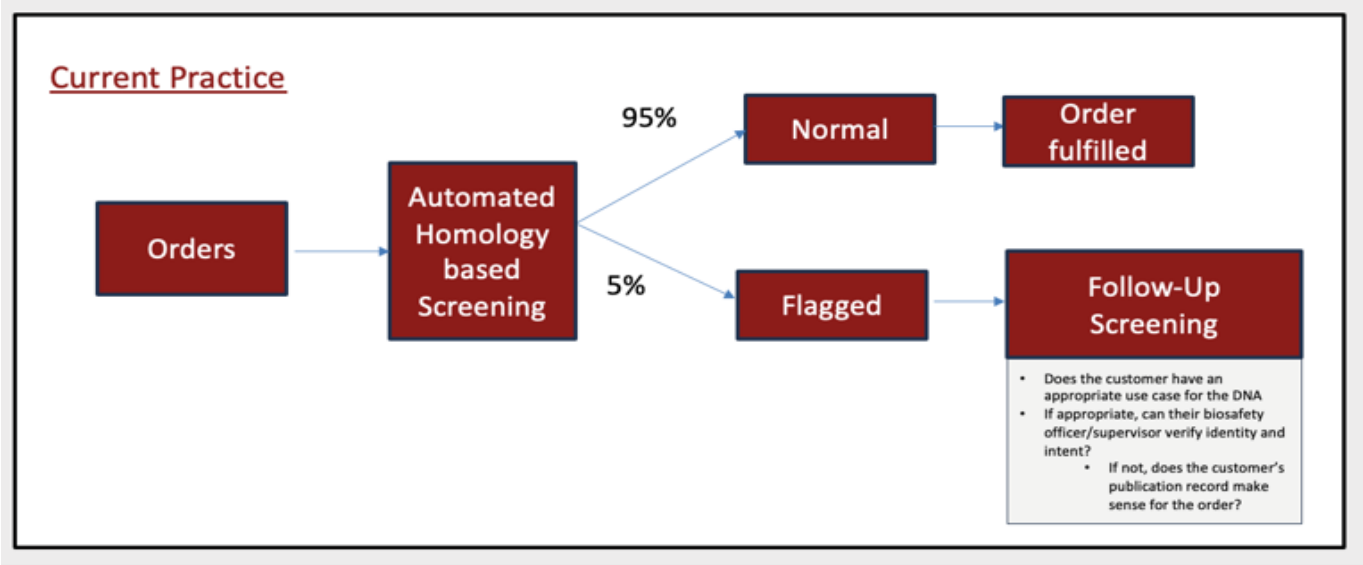

Figure 3: Current screening practice flag and manual review rates

As highlighted by Figure 4, on a fundamental policy level, it is unclear who the policy owner for VWMD prevention is, and there are a number of overlapping roles that various groups hold within the field. This makes it challenging for researchers and biosynthesis firms alike to understand where to receive information and support for biosecurity compliance efforts.

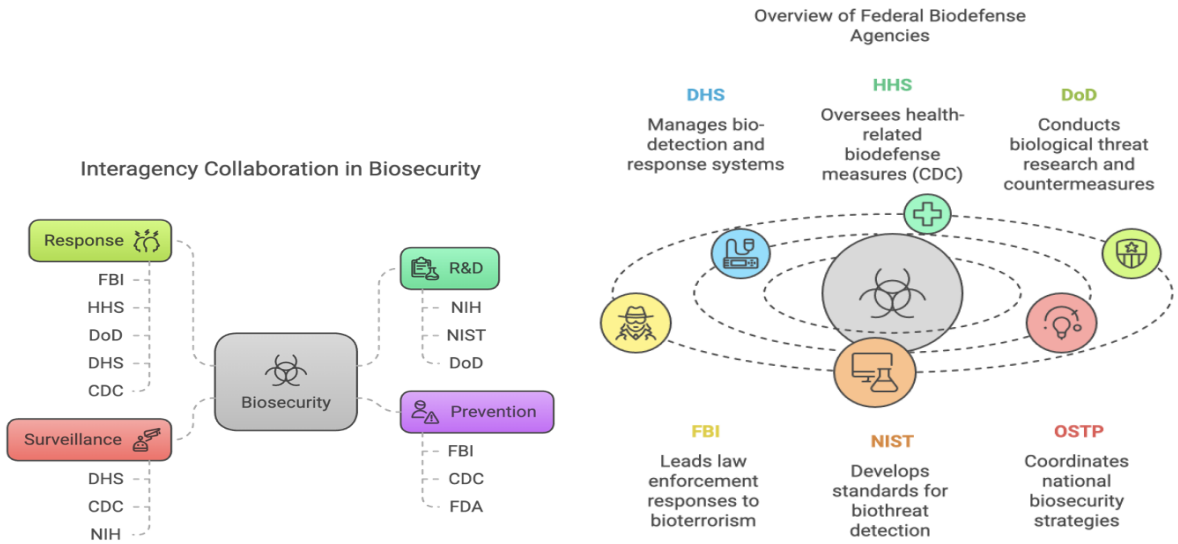

Figure 4: Cross-agency roles and coordination of VWMD biosecurity responsibilities

The most active role held in the VWMD prevention field is carried out by the FBI. The FBI established the Synthetic Biology/Emergent Biotechnology initiative in 2011, which is a

proactive approach to mitigate current and future horizon risks posed by the exploitation of synthetic biology and nanobiotechnology techniques (15). Overall, the FBI functions as an enforcement agency and is responsible for more acute issues; notifying of critical incidents through Weapons of Mass Destruction Coordinators in each state.

The CDC engages in a number of bio-surveillance efforts that can help identify abnormal viruses for escalation for intelligence support.

The Office of Science and Technology Policy (OSTP) coordinates across agencies, but its role is generally to facilitate consensus, set research and development priorities, and oversee framework development, rather than to directly enforce policy. That coordination remit was blurred via the Framework for Nucleic Acid Synthesis Screening, which was issued on April 29, 2024. The framework set out provider obligations, requiring procurement of synthetic nucleic acids only from providers that attested to screening standards, which created a binding condition for NIH-funded work.

The NIST sets documentary standards, reference materials, and measurement protocols for biothreat detection technologies, but its standards remain voluntary for private-sector companies, leading to inconsistent adoption

Due to only voluntary requirements for most customers, only large companies undertake screening for customers, largely because they can afford to dedicate resources to screening. Many of these large companies (about 80%) are part of the International Gene Synthesis Screening Consortium (IGSC) (39). These IGSC member companies have agreed to adopt DNA synthesis screening protocols. However, this voluntary nature with minimal enforcement fails to address the fact that a VWMD can be constructed anywhere and spread to harm communities globally (44). National bodies such as the FBI are limited in their ability to stop such vast threats from both a resourcing and authority/jurisdiction perspective.

Overall, this landscape highlights a gap in coordination over enforceable biothreat prevention standards. While multiple agencies contribute vital capabilities from surveillance to enforcement, the absence of a clear authority to mandate and monitor biosecurity compliance across both public and private industries impairs action. The closest currently deployed enforceable framework is the Federal Select Agent Program (FSAP), which regulates possession, use, and transfer of a small subset of listed agents. FSAP is jointly managed by CDC's Division of Regulatory Science and Compliance and USDA's APHIS Division of Agricultural Select Agents and Toxins.

The 2024 EO was set to be in force April 29, 2025, but was deferred while revisions proceeded. The Executive Order on Improving the Safety and Security of Biological Research (May 5, 2025) acknowledges that the previous EO did not cover the whole market and seeks a path to mandatory and verifiable screening (45). It directs OSTP and partners to develop a strategy for comprehensive, scalable, and verifiable screening in non-federally funded settings and to prepare a legislative proposal.

Cost per screen and complexity of balancing approaches for pragmatic policy

The 2025 EO establishes a strong call to action for increased screening compliance. However, as agencies begin to draft and execute on its mandates, they must consider barriers in the current landscape and the downstream consequences of policy design options.

The fundamental challenge with screening adoption is that producing synthetic DNA **costs cents,** whereas current screening systems (when factoring in manual flagged order review time for personnel), **cost in the magnitude of dollars**.

A rough threshold that screening must cost is less than one dollar, to ensure synthesis production is financially sustainable. This highlights why adoption is fragmented and private acquisition of biosecurity tools presents systematic biases against smaller firms.

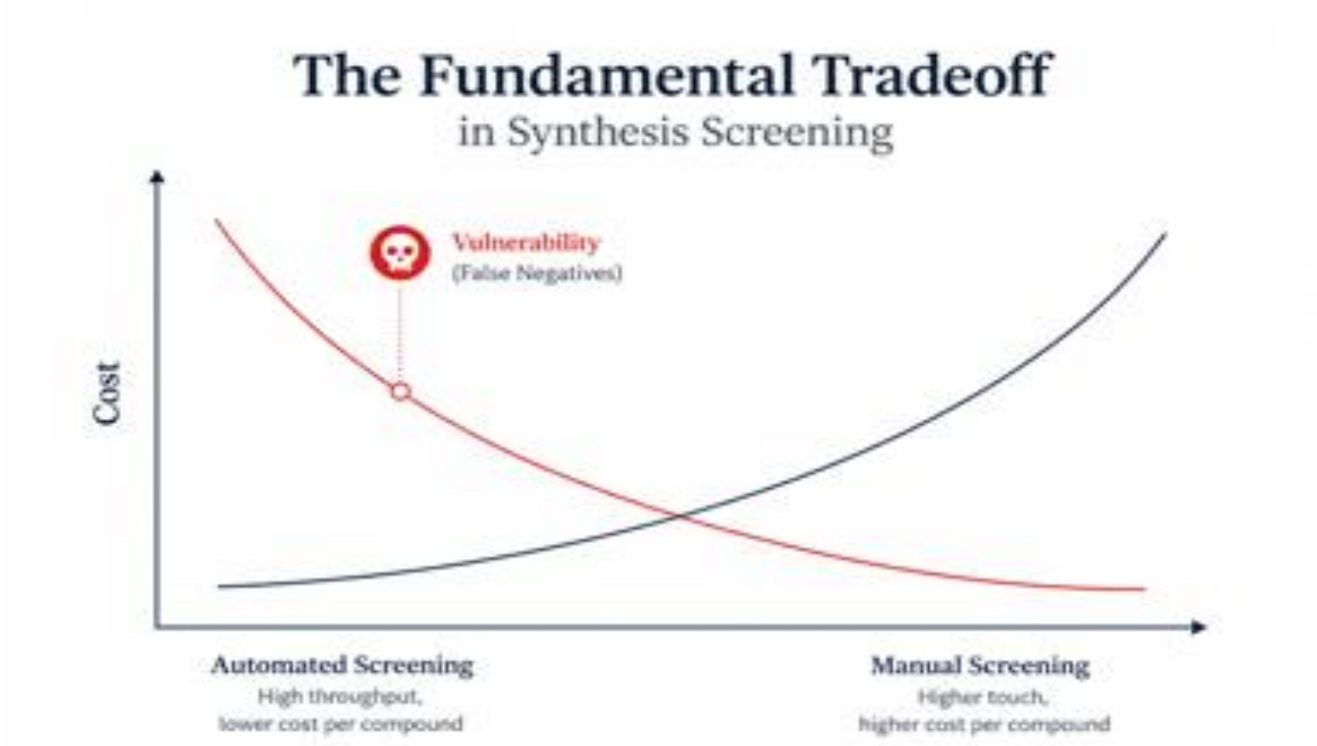

Figure 5: Visual highlighting the trade-offs between high automated screening with greater false negatives (less secure) but lower cost, vs manual systems with low false negatives (more secure)

For new startups or small to medium-sized enterprises, there is less funding to hire dedicated biosecurity personnel, and/or institutional experience/awareness of such biosecurity issues, thus creating a gap for bad actors to exploit. This cost factor is a critical gap that is challenging to address without sufficient market incentives or financial support. Although free, open-source screening tools may decrease costs, it is fundamentally this human resource cost factor that impedes implementation.

Figure 6 highlights the practical challenges of system design via presenting various screening modes, drawing upon prior studies examining the landscape (42,43). These design considerations are crucial for policymakers to create pragmatic policies that can enable the synthesis industry to thrive while balancing security concerns. Particularly, the implications of increased manual review flags due to larger market adoption and the increased sophistication of protein design tools, or the cost burden of adding manual order review stages to screen all orders.

This analysis of design modes also invites exploration of acceptable specificity and sensitivity thresholds. Existing policy frameworks do not specify performance metrics in terms of overall sensitivity and specificity, or the magnitude to which sequence perturbations or annotation sources acceptably should impact screening results (46). On one hand, stipulating requirements may encourage higher-performance tools and help emerging firms have a rough idea of acceptable quality, thereby reducing gaps. However, similarly, such requirements may also create cost barriers that impede the financial viability of biosynthesis business models.

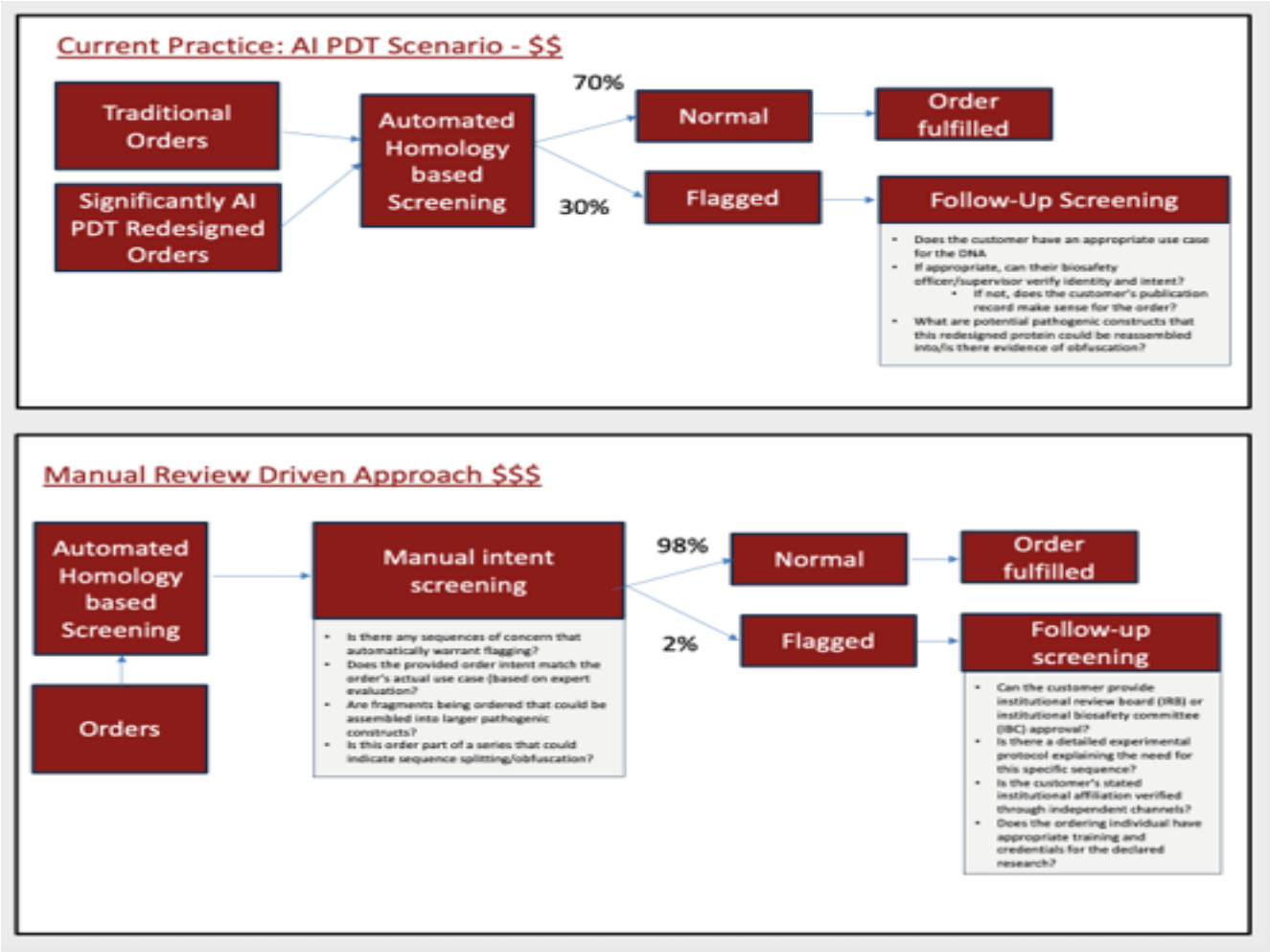

Figure 6: Overview of various financially unsustainable screening modes. Box 1 highlights a future where a high prevalence of AI redesigned proteins causes high manual flag rates, and Box 2 demonstrates a redesigned manual review-centred screening process

Defining biothreats and an unclear understanding of risk.

Companies voluntarily engage with weapons of mass destruction coordinators from the FBI to report suspicious orders. On an individual level, these relationships act as valuable resources for support and advice. However, as this is an emerging field, there are no guidelines for what is suspicious, and companies must largely exercise their own judgment. It was reported in one study that less than 50% of interviewed personnel knew they could report concerns to the FBI, and less than a quarter could say with certainty that they knew the company protocols for reporting (43).

Put bluntly, it is simple to decide to reject an order for Ebola – it is the more borderline cases that present difficulties and require further reflection. Consider the challenge of defining risk for the following cases, A*n influenza strain from a repeat customer who has a new PO box shipping address. An order for a new customer for a pathogenic virus that is not on the federal select agents list.*
A pathogenic virus order for an early career virologist, but the laboratory does not have the facilities to carry out their stated research objectives.

Standardized case studies or examples of bad actors using PO boxes, masked ordering techniques, or making chimeric proteins are needed to help provide clarity on such cases. Private AI frontier corporations such as Anthropic and OpenAI, among 50 other signatories spanning academic and national security domains, have recently signed a public letter to the US Congress to make screening of orders for synthetic nucleic acids and related

equipment, mandatory. The letter notes that *"there is a real possibility that the knowledge barriers which have historically prevented bad actors from obtaining biological weapons will meaningfully erode"* (47). The authors urge to policymakers that *"to ensure a consistent national standard rather than a patchwork of conflicting laws, states should also consider implementing requirements based on existing federal and industry guidelines"* (47).

Understanding risk will be crucial for compliance with the new federal screening provisions under the Executive Order on *IMPROVING THE SAFETY AND SECURITY OF BIOLOGICAL RESEARCH (*45*).* The EO calls for effective encouragement of providers of synthetic nucleic acid sequences to implement comprehensive, scalable, and verifiable synthetic nucleic acid procurement screening mechanisms to minimize the risk of misuse. While this is a step in the right direction, it must incentivize improved biosecurity measures without friction, or risk negatively disrupting market forces in the bioeconomy.

An action/implementation-oriented framework must be created to define key actions for those overseeing federally funded research, and an extension to private industries and non-federally funded research thereafter is critical. As it will be explained further on, what constitutes as a "verifiable' flag or "comprehensive" screening process, will be challenging to define.

Current Gaps in VWMD Prevention

As previously analyzed, the current threat level for VWMDs is low. However, a confluence of biosecurity gaps could lead to the development of a VWMD from bad actors. Visualizing the various layers that gaps lie, can be best done via a Swiss cheese model (48).

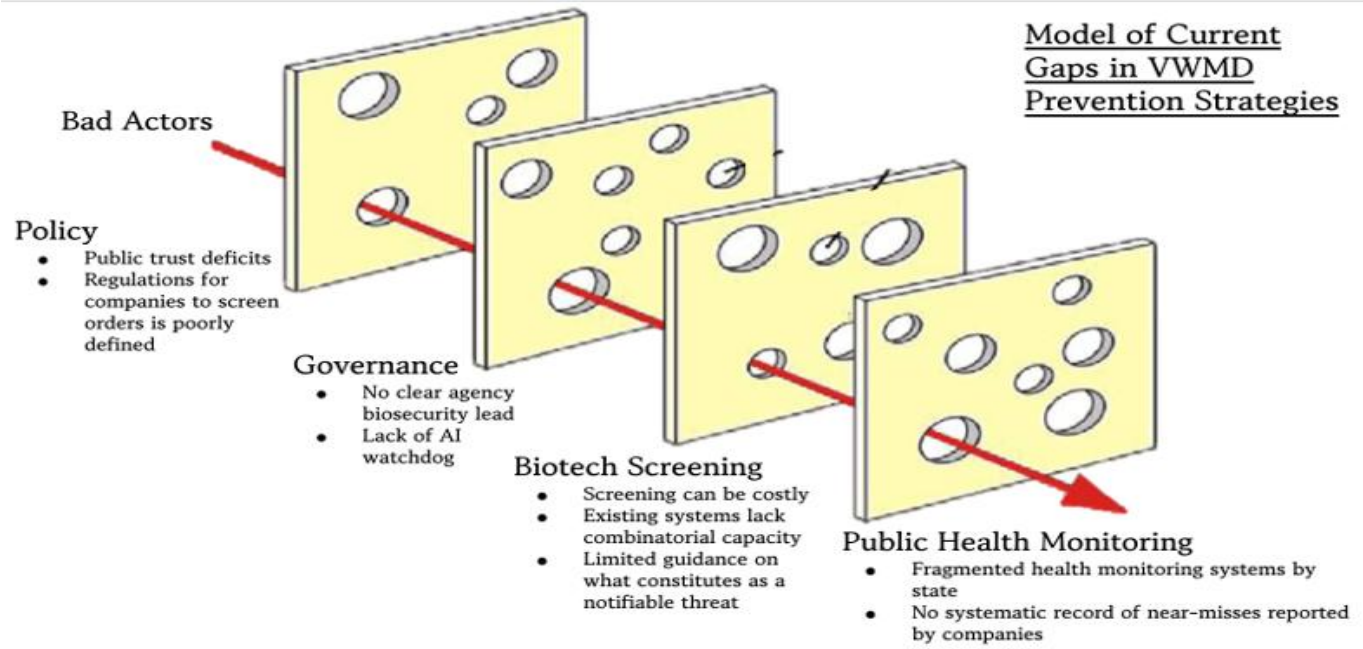

Figure 7: Diagram highlighting the current incapability of screening systems to detect combinatorial orders, adapted from J. T Reason (48)

Sequence size limitations and Combinatorial ordering

A significant gap/pathway for circumventing synthesis screening is via combinatorial ordering. Threat sequences can be divided across multiple orders, with individual fragments falling below screening thresholds. While companies advertise fragments as small as 100-125 base pairs, their capacities are developing rapidly with Gibson assembly enabling 15-20 BP overlaps and Twist Biosciences offering 20 base pair oligonucleotide synthesis. However, most screening tools can only capture up to 30 base pairs, even though oligonucleotides can range from 13-300 base pairs (49). This provides a crucial gap for fragments that fall below this threshold.

The EBRC's guidance only suggests baseline standards for screening sequences using a "best match" system in which each 200-base pair span of an ordered sequence is matched (39).

Beyond screening of these fragments ordered from the same company, current screening systems completely lack detection of distributed ordering due to a lack of customer data sharing infrastructure (49). This enables bad actors to order separate fragments across different providers and reassemble a complete pathogen, without detection.

While it is argued that base pairs below this length of 30 are unlikely to be able to be feasibly joined together to create stable viruses, the risks are not zero and can change at any moment as technologies develop (50,51). Reliably modelling for threats also becomes near impossible at smaller base pair levels as the folding combinations for these proteins become countless. Alternatively, adopting a more risk-averse approach to risk triage of small BP length risks high false-positive rates. This would drive up manual reviews and make screening unsustainable, thus creating pressure to reduce sensitivity and potentially allow dangerous sequences to pass undetected.

Moreover, it has been reported that common housekeeping genes (non-pathogenic regulatory components) from pathogens can trigger false alarms, leading to screening fatigue (42).

Reassuringly, the current administration has mentioned the need for distributed order data sharing in the White House's "America's AI Action Plan (52). Now that this is on an official agenda, work can be done on proposed data sharing arrangements and data custodianship to evaluate the feasibility of this proposition. As will be discussed later, there are many challenges with establishing data sharing due to competing industry incentives, technical implementation feasibility, and intellectual property concerns

Incomplete Pathogen datasets

Many companies follow the NCBI Pathogen Detection Portal's core pathogen catalogs and the federal Select Agents and Toxins list (SAT) under the Federal Select Agent Program (FSAP) to identify sequences of concern (39).

These lists are not all-inclusive and feature significant omissions, such as "functionally equivalent" or engineered pathogens that are not taxonomically on the select-agent roster (53). For example, viruses reconstructed by swapping antigenic or polymerase domains, Sub-genomic elements, and virulence-associated sequences, such as toxin subunits (botulinum heavy-chain fragments, anthrax protective antigen, staphylococcal enterotoxin domains) (53). Emerging and re-emerging pathogens identified after the last FSAP update, most notably novel filoviruses, which bear similarities to Ebola (BOMV, LLOV, MLAV) are other omissions.

Even within the FSAP list, portions of a select agent genome could be ordered without being subject to FSAP regulations and later assembled and recovered to form the original agent or combined to enhance the virality of another pathogen, creating a chimeric agent (53).

US screening providers also adhere to country-specific export lists for international orders. Similarly, they are list-based systems that encompass a limited set of pathogens that could also be unrepresented in FSAP. These export lists constantly change and have limited overlap in some contexts. This presents an issue for international collaboration wherein some countries may perceive certain viruses to be threats vs not for others. At the very least, for future international data-sharing approaches for threats, there should be some alignment on a core list of pathogens, and a sort of extended list for each country.

Architecture Limitations of Current Computational and AI Biodefense Tools

Last year, there was a zero-day event across several screening providers, where researchers from IBBIS, RTX BBN, and Microsoft discovered that AI-powered protein design tools were able to create synthetic toxic protein variants that bypassed existing DNA synthesis screening tools. This led to a year-long project involving other screening providers to patch their tools, culminating in algorithms that could detect synthetic homologs at 97% efficacy (12). While no adverse consequences have been reported as a result of this vulnerability, it reinforces the importance of pre-emptive adversarial testing as protein design tools evolve. The authors noted that even after patching, across providers about 3% of the variants, which were most likely to retain functionality, still escaped detection (12). This highlights that despite best efforts, there will remain some missed threats, and that

homology-based screening architectures are fundamentally flawed.

Homology-based screening presumes that pathogenic function is always correlated with sequence similarity to known threats. However, in biology, this is fundamentally flawed as different proteins can be used to create the same function, and sequences can be altered through changing primers, non-coding sections, or duplicating sections to sneak past detection across such systems (42,43,54). Substitution attacks are a threat vector that bypasses many current screening safeguards (11). These attacks exploit synthesis instrument vulnerabilities by physically swapping nucleotides during the synthesis process, enabling the creation of dangerous sequences from apparently benign orders (11).

Recently, researchers created an LLM agent that used existing protein design tools to convert non-pathogenic protein designs into pathogen-like sequences that evaded detection by several screening methods (up to 60% Attack Success Rate for Evo2-40B) (55).

Functional homology-based screening

A proposed pathway for addressing these current limitations is through functional homology. Functional homology–based screening seeks to go beyond simple sequence-similarity alerts by identifying whether a specified DNA sequence would likely fold into a protein structure that could bind to produce pathogenic functions. EVO2, AlphaFold, and other tools are already being used by synthesis firm companies to interrogate flagged suspicious orders, essentially as a second-line screening tool. However, while such tools are powerful, there are inherent limitations that limit their use for frontline screening.

Generating a full 3D model for each candidate protein, even using GPU-accelerated AlphaFold-3, involves minutes of GPU time per sequence. If synthesis firms were to screen at scale, the total compute cost could exceed millions of dollars per year. This is easily incompatible, considering the profit margins on synthesis orders are so low.

An alternative, more scalable strategy in development seeks to leverage machine learning and protein sequence embeddings, low-dimensional vector representations that capture both local sequence patterns and remote homology signals, to proxy for structural and functional similarity without 3D modeling (56).

Ultimately, there are a range of technical limitations that still exist for functional homology-based approaches, such as false positive rates stemming from incomplete data and an inability to even accurately predict certain structures. Structure-based classifiers require large volumes of human-annotated training data for each fold family of interest and exhibit reduced sensitivity on small, disordered, or multi-domain proteins.

A promising collaboration between OpenAI and Lawrence Livermore National Laboratory, is helping improve preparedness, accelerate response, and strengthen confidence in promising countermeasure development. The decision for OpenAI to launch Rosalind Biodefense to help trusted developers to build new biodefense and pandemic preparedness capabilities, is a welcome move that will aid private industry to resolve this gap left by Congress. Rosalind is reportedly supporting specialized companies such as Fourth Eon Biosecurity, to build adaptive screening that evolves alongside next generation frontier models to achieve reliable function-based screening that can handle completely de novo proteins.

Specific technical limitations of current function-based screening technologies (functional homology)

Modeling protein-protein interaction (PPI), particularly epistatic couplings in which different sites cooperatively influence each other, is a key weakness (57). The accuracy of binary PPI prediction itself is contested, requiring careful consideration of data leakage and contamination in evaluation approaches (58). Accounting for solvent effects is also a crucial but often overlooked aspect of protein function prediction (59).

Current approaches suffer from an overemphasis on global sequence similarity and nodes, rather than the nuanced protein pockets and domains crucial for function (60). Understanding multichain dynamics and allosteric changes is hugely important but poorly integrated in current models (61,62). In nature, pathogens often use induced-fit binding to optimize evasion of host defenses which is a complex dynamic biophysical process that current tools, which essentially look for functional mimicry, are unable to handle (63).

Function-embedding-based tools are vulnerable to sequence perturbations; merely changing several surface residues can drastically reduce sequence identity while maintaining binding affinity. Additionally, this can go so far as codon swaps, or entirely synthetic codons as seen in the work of the Chin lab (64). Synonymous substitution, using genetic code redundancy to change sequence similarity, can bypass automated systems. Overall, through utilization of entirely synthetic codons, or structural persistence engineering, an adversarial AI agent could create a protein that folds perfectly into a pathogenic structure, yet whose embedding vector is mathematically orthogonal to any known threat cluster within the screening tool's latent space

Another fundamental challenge in building these embedding tools based on existing data is that our data is based on the constraints of evolution, rather than understanding the possibilities of viral design, which encompass uncharted regions of sequence space. This creates "blind spots" in areas where novel bioweapon threats may emerge, as models overfit to training templates and evolutionarily generated variants, vulnerabilities that adversarial actors could exploit. Thus, there is a critical need for validation of truly *de novo* virus designs, as opposed to just evolution-inspired variants, to assess real-world resilience. However, as will be described later, validating pathogen designs in the real world is an inherently ethically fraught task.

Despite efforts to create the "best AI models," there is insufficient work on making manually annotated functional training and testing datasets within virology. As large as 86% of environmental viral protein clusters match uncharacterized families of proteins or have zero data (65). Thus, no matter how much these architectures advance, they must overcome a foundational labeled data gap. Manually annotating this data is highly costly, and AI approaches are limited in their capacities to annotate critical domains (65,66).

Ultimately, this is not to say that functional embedding tools are useless; they serve a valuable role in “raising the floor”, capturing less resourced bad actors while discouraging others from pursuing other means of acquiring weapons for terrorism. However, the practical challenges of deploying current functional embedding-based architectures into screening require strong consideration of potential limitations.

Critical Gaps in Biosecurity Governance and Policy

Current governance structures present significant hurdles to effective biodefense and the safe advancement of synthetic biology.

The challenge of Triage

A critical issue with the new EO is verifying legitimate threats from potential threats and utilizing a comprehensive system, which also does not render standard business unfeasible. As mentioned previously, the critical threshold for screening is $1 per molecule. This number is an average figure, which is calculated by a cumulative total of screening software costs + human labor in manually flagging, investigating, deliberating, and taking the final action, whether it is to ignore or report. Companies report that current screening systems are reaching this cost.

Currently, screening algorithms automatically classify compounds into several categories of risk. A comprehensive manual screen at the extreme end is what is conventionally reserved for flagged suspicious orders (unflagged orders simply pass through uninterrupted). Comprehensive manual screening involves analysis of order pathogenic potential (based on 3D reconstruction and classification of type + function of viral protein, similarity to existing pathogen proteins, legitimate use cases analysis and a check for listing on federal select agents programs) and a follow up screen, checking for customer intent (request for information on proposed use case, shipping address suitability, publication track record, affiliated university/institution) (39). Many of these check components are not applied uniformly and to varying degrees as there are no universal enforceable standards. Overall, performing a comprehensive check for every single DNA molecule ordered would be unfeasible; thus, automated systems for initial screening are crucial.

Thus, the trade-off for what criteria orders should meet to justify this manual review/how to appropriately and sustainably allocate human resources to biosecurity efforts, is a new arena. One such challenge is the advent of AI-edited strains, which have no homology matches in screening databases. These strains may have slight changes to their sequence coding that may have drastic changes when folded in 3D space. Thus, simple percentage-

based matching for similarity focused on list-based homology alone is insufficient, and the need to manually flag each of these leads to raised costs. It has been reported by some synthesis firms that up to 10% of currently screened orders have been AI-edited, indicating that this is already a horizon 2 issue.

However, screening platforms have now embedded algorithmic changes that account for some level AI-design, thus not all of these AI-edited proteins are proceeding to manual review. As AI protein design tools advance, the creative possibilities will likely exceed the capacity for screening tools to predict function using current architecture. As will be discussed later, robust 3D folding-based/function-based screening would be extremely challenging to create and deploy with cost efficiency.

The verification of a potential threat to a real threat also raises legal concerns. For enforcement agencies to execute search warrants or use intelligence software against bad agents, a necessary verifiable suspicion level is necessary. The level of evidence needed to support such actions must be defined to ensure that flagged orders are even able to be acted upon/worth investigating. In cases that accidental or hasty actions occur, undue reputational harm or work delay may occur. In the case that screening software/synthetic provider personnel are unable to justify their concerns, and flags are acted upon, they risk invading the privacy of legitimate researchers trying to do their job.

Combinatorial orders and know your customer

A critical current gap is that individuals could make nucleic material orders across several companies, known as combinatorial ordering, where they can order pieces that individually don't set alarms, but can be combined to produce bioweapons. This is a well-known fact in the narcotics field, "You don't buy it all from the same place". The EO on improving nucleic acid screening seeks to address this gap with distributed order tracking; however there has been limited description on how this would be deployed *(45)*.

Know your customer laws have been highly beneficial in the financial sector and is a critical prerequisite for any distributed order tracking system. However, for the bioeconomy, there is a lack of systems to verify the intent and identity of buyers, and to track the flow of material post-purchase to monitor the end users of synthesized compounds. Companies currently collect administrative information and use order addresses to provide a base level of assurance. In the past, some companies have tried to integrate standard questions in order forms to understand what customers intend to use products for.

However, customers report that filling out such intent forms is cumbersome, and this risks causing them simply to move on to the next provider. Thus, it must be a whole of industry deployed standard for such mechanisms to work. Moreover, the human workforce impact of needing to manually review intended project objectives and screen compounds for intention/potential use concordance could add unfeasible personnel cost. If instead LLMs/AI systems are used to lower cost, there is then a risk that such systems could be gamed or have gaps.

Establishing incentives to bolster biosecurity measure adherence

Screening companies could enhance uptake by adopting multifaceted value propositions. One current offering by IBBIS is locale export ban screening, enabling synthesis companies to comply with restricted compounds lists, which drastically vary by country. Thus, it removes the pain point of companies having to manually maintain export ban screening protocols for every country.

A virologist' "clearance" or "Passport", a background check administered either by a private organization or government, could enable screened individuals to have fast-tracked orders. Such a system could be a worthwhile incentive as it could reduce administrative delays/order vetting requirements for qualified personnel. From the public safety side, this model could allow for stronger governance and insight into private sector virology research. However, the extent of information that would need to be provided to establish a robust system requires a balanced approach to ease privacy concerns. Similar schemes have been successfully deployed by the Global Alliance for Genomics and Health for genomics data access management (16). Reportedly, SecureDNA is developing tools to help screen for customer qualifications, cryptographically secure certification methods to examine if individual researchers have been approved for work with various biological agents (16). However, such

methods fail to address biohacking or DIY threats, as material may be shared within these groups with limited federal oversight or governance.

Companies can find it challenging to understand what sequences to report, vs not, but understand that there are huge liability risks for poor risk management. Having a risk triage system that automatically assesses risk factor information and categorizes threat levels can help encourage adoption. Such a system could rewrite the cost/benefit perception and make implementation less of a resource burden. Developing standardized templates for FBI reporting would not only reduce deliberation time for synthesis firms but also help the FBI in executing appropriate threat responses. Such templates could encompass what order information is suspicious, are there are inconsistencies in use cases for ordered agents vs provided rationales, and request information on the professional background of the customer. For extra usability for enforcement, it would be valuable to have a percentage accuracy or confidence in each of the risk factors.

Several ideas have been proposed to improve biosecurity for the synthesis industry. However, evaluating the feasibility of each is out of scope for this paper. Nonetheless, Figure 8 highlights existing mechanisms and potential options to increase awareness.

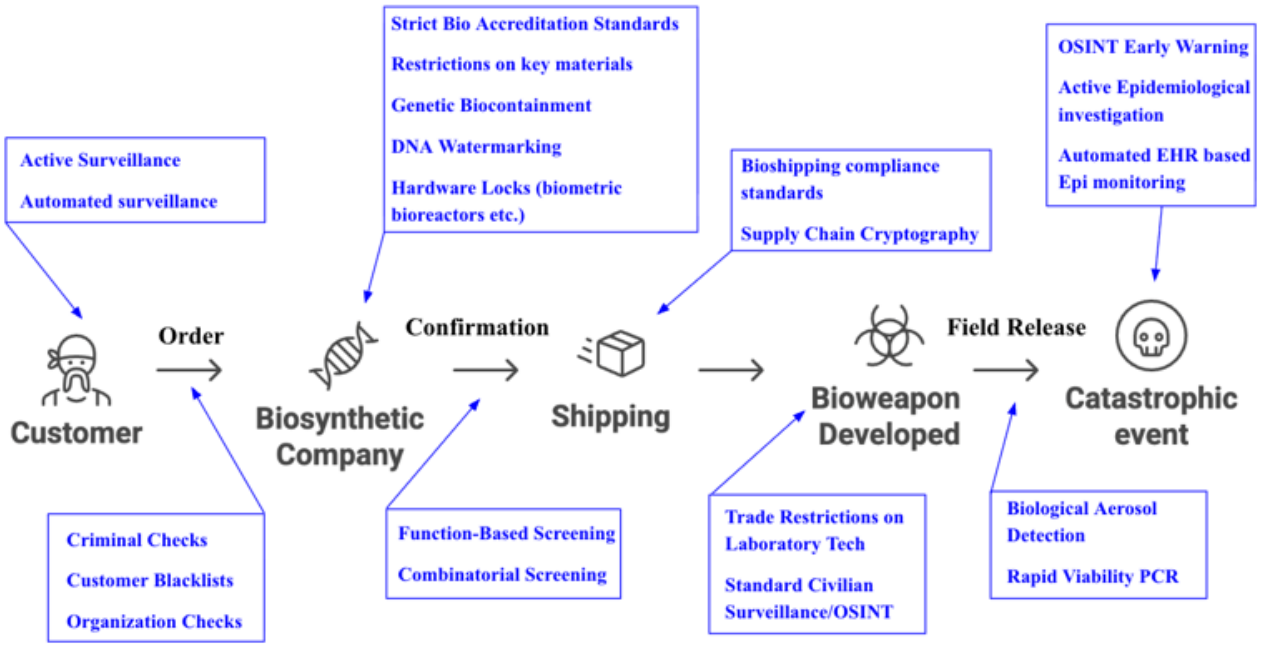

Figure 8: Visual overview of existing and potential solutions for improving biosecurity for the synthesis industry

Ethical and Regulatory Dilemmas:

A profound ethical dilemma is conducting research to confirm whether AI-designed pathogens are in fact viable/pathogenic. Translating designs from code to actual physical proteins would directly violate international treaties such as the Geneva Protocol (1925) and the Biological Weapons Convention (BWC, 1972) (20,43). Furthermore, publicly or government-funded research conducting such tests might inadvertently give adversaries the impression that biological weapons production is underway, leading to an unintended arms race.

Another regulatory issue is managing corporate control over trademarked/proprietary data vs the need for data sharing for biosecurity measures. The synthetic biology products being ordered from synthetic biology firms are largely proprietary. Pharma firms already find it challenging to share orders with external companies, let alone adding another vendor to be privy in the form of screening companies. Such data can be incredibly valuable for competitors. Leakage of this data could enable replication of intellectual property, such as proprietary formulations or structures, and knowledge on order volumes could enable

diversification and capture of market share by competing companies. Privacy over order contents may also drive opposition to government involvement in synthesis screening or order logging. This selection pressure from synthetic firm clients pushes screening companies to integrate confidentiality into tools.

Creatively, SecureDNA has integrated a cryptography system where neither the screening company nor the biosecurity officer is privy to the contents of the order (49). Flagged orders necessitate closer evaluation and exposure of the order's contents; however, this is reasonable given that only a small fraction of orders are flagged and reviewed.

A major concern is the lack of public oversight for private companies developing these technologies, leading to poor governmental control and visibility. This creates a dangerous opacity where critical safeguards might be overlooked. So far screening has been encouraged by the government yet driven by industry (private screening tool providers), which inherently presents risks. A lack of public investment in screening tools and screening mandates has caused market forces to define solutions. These center around preventing reputational damage associated with firms accidentally supplying pathogenic materials to bad actors.

This environment has enabled sustained innovation of currently available tools (originally born largely out of academia). However, instead of enabling market selection pressures to select for workflow suitability (thereby primarily serving the interests of companies), it would be better to remove this lever and design tools based on public agendas and adherence to (one day) well established federal screening guidelines. However, achieving this requires large public investment, and clear federal guidance and commitments to support long-term maintenance of such systems.

To demonstrate how such a system could work, and where biosecurity investments should ideally be targeted towards, we have proposed a strategic biosafety flywheel.

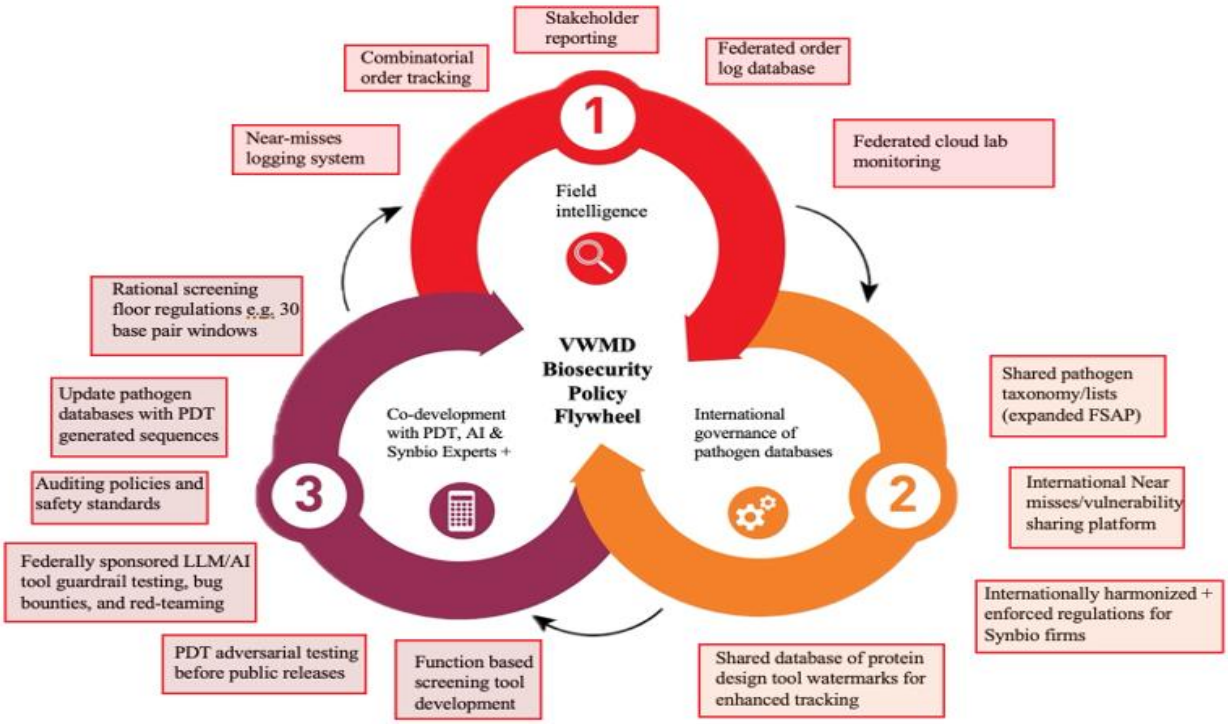

Figure 9: The biosecurity policy flywheel for VWMDs

The VWMD Biosecurity Policy Flywheel

The flywheel is centred around leveraging current screening tools, promoting cross-sectoral collaboration, and pooling resources to help advance function-based screening (as investing in such blue-sky technologies requires long-term and patient capital).

Existing screening tools can benefit from being integrated into federated systems that can track combinatorial orders or repeated flagged/borderline risky orders across biosynthesis firms. This can leverage the existing recommendations from the recent EO, to maintain records of orders; however, it would require significant interoperability and data linkage efforts. It would also be necessary to extend order retention from 3 years to indefinite time spans to account for the long lead-time and iterative nature of VWMD development (40). Preventing future zero-day events requires collaborative efforts between protein design tool (PDT) developers and synthesis screening tool providers to ensure tools can handle increased PDT capacities before they are publicly released.

Similarly, a near-miss program following an interdisciplinary community of practice model could be highly valuable. Such a community could focus on sharing deidentified case studies of real pathogenic orders that were nearly shipped out to bad actors. As AI tools become more sophisticated at obfuscation, it is vital to share such learnings across the biosafety ecosystem (53).

A barrier to achieving an effective functional screening tool is the lack of labelled viral data. By pooling investment across nations, this could fund manual annotators and real-world lab testing of analogous non-human transmissible AI edited viruses as proof of concepts/training data to help AI models learn. The goal would be to understand the limits of genetic drift whilst preserving functionality for key pathogenic subunits. Such a rich dataset could also be used by nations to develop vaccine targets and enhance broader scientific understanding of viral pathogens. Balancing usability with minimizing vulnerability to penetration by bad actors is important. Cryptography could be a part of the solution (as discussed previously with SecureDNA's approach), as this could also help protect commercial interests surrounding proprietary sequence order information.

While it is sometimes argued that more manual review and know your customer efforts are needed, the practicalities of implementing practical, defensible systems revealed unsuitability as discussed earlier. The fundamental issue is that screening molecules can cost upwards of dollars due to the personnel time required for manually reviewing orders.

Even currently deployed hybrid systems are constantly being tuned to drive down false positive flags to ensure screening does not reach the critical threshold of one dollar per screened molecule, where it becomes unsustainable to conduct business. This is especially an issue in an age where protein design tools are constantly evolving, and significantly AI-designed orders can cause increased flagged orders for automated screening tools. Thus, the financial viability of operations is liable to flag rates, and the equation could change at any time. This drives home the importance of collaborative efforts between protein design tool firms and synthesis screening providers in reducing the onus on bioinformatics professionals at synthesis companies to tune their screening systems.

International Cooperation and Data Governance

Achieving international cooperation on such a divided issue is challenging but necessary in an increasingly interconnected world. If a sequences of concern database lacks listings of viral variants being designed in one corner of the world, it systematically undermines the efficacy of all other models when such viruses spread across borders.

Furthermore, the ability for bad actors to simply run AI searches and quickly identify other synthetic material providers that don't screen internationally nullifies the efforts of domestic institutions to enact screening (49). International cooperation to regulate access to emerging high-risk technologies such as benchtop synthesizers will further become of paramount importance.

An internationally maintained database of sequences of concern will help enable collaboration and universal understanding of concerning sequences. This can be further aided by a comprehensive record of watermarks that different international cloud lab providers and protein design tool firms to help tracking. As international

synthesis industries grow and their firms become market leaders, it's possible that their compliance with order records management will become a bottleneck for global biosecurity risk prevention efforts.

The NTI has highlighted this global vulnerability in biothreat prevention and called for intelligence sharing mechanisms and efforts to analyze social media, trade, published research, and genetic sequences to detect potential biothreats before they occur (67). They have further recommended that the UN create a Response Coordination Unit to prepare for and coordinate complex responses to high-consequence biological events. It is clear that such a specialized international unit is necessary given the level of governance and resourcing needed to analyze multinational data for agentic R&D patterns, monitor the traffic of global genomic sequences, share best practices post responses to zero day events, and coordinate rapid responses to high consequence biological events driven by the next paradigm of development of Artificial intelligence.

Conclusion

The current path of synthetic biology and AI development holds much promise for society; however, it also brings about serious biosecurity risks that existing tools and governance cannot keep up with. Current tools fail to account for the risks posed by PDTs, and a lack of mandated screening in existing policies has resulted in fragmented biosecurity standards adoption. To move forward in addressing these challenges and to promote responsible innovation in an agentic AI era, we need strong public oversight and investment in effective policies that prioritize public safety over commercial gain.

The risks that current technologies pose are understood, yet there is insufficient development of resources to assist prevention efforts. Currently, there is a lack of reasonable screening tools that make sense for companies to adopt due to cost, explainability, and workflow integration issues. We need joint international collaboration and investment in tools to coordinate pragmatic, harmonized efforts. Integrating global standards could include improved transparency and enhanced global protection, and for companies, they could help them scale drastically across the globe through eased regulatory hurdles.

There is work underway to establish a synthesis screening consortium, an industry-led initiative involving screening companies such as IBBIS, Raytheon BBN, SecureDNA, Aclid, and Battele, and synthesis providers such as Twist Biosciences, to address some of these gaps. Ultimately, effective approaches require comprehensive cross-sectoral engagement from policymakers, PDT firms, biosynthesis companies, and screening providers alike. The future success of biosecurity policies will depend on the extent of cross-sectoral collaboration and incentives alignment.

References

1. Bell JA, Nuzzo JB. Global Health Security Index: Advancing Collective Action and Accountability Amid Global Crisis. Nuclear Threat Initiative and Johns Hopkins Center for Health Security; 2021.

2. Chen X, Kunasekaran MP, Hutchinson D, Stone H, Zhang T, Aagerup J, et al. Enhanced EPIRISK tool for rapid epidemic risk analysis. Public Health. 2023;224:159-68. https://doi.org/10.1016/j.puhe.2023.08.032

3. Jefferson C, Lentzos F, Marris C. Synthetic biology and biosecurity: challenging the "myths". Front Public Health. 2014;2:115. https://doi.org/10.3389/fpubh.2014.00115

4. Franco D, Pathak HB, Cameron CE, Rombaut B, Wimmer E, Paul AV. Stimulation of poliovirus synthesis in a HeLa cell-free in vitro translation-RNA replication system by viral protein 3CDpro. J Virol. 2005;79(10):6358-67. https://doi.org/10.1128/JVI.79.10.6358-6367.2005

5. Imperiale MJ. Re-creation of Horsepox Virus. mSphere. 2018;3(2). https://doi.org/10.1128/mSphere.00079-18

6. Newswire. HS. How secure is gene synthesizing research? Homeland Security Newswire,. 2024 2024-06-08 06/08.

7. Adamala KP, Agashe D, Belkaid Y, Bittencourt DMdC, Cai Y, Chang MW, et al. Confronting risks of mirror life. Science.

2024;386(6728):1351-3. https://doi.org/10.1126/science.ads9158

8. King SH, Driscoll CL, Li DB, Guo D, Merchant AT, Brixi G, et al. Generative design of novel bacteriophages with genome language models. bioRxiv. 2025:2025.09.12.675911.

9. O'Brien JT, Nelson C. Assessing the Risks Posed by the Convergence of Artificial Intelligence and Biotechnology. Health Secur. 2020;18(3):219-27. https://doi.org/10.1089/hs.2019.0122

10. Arias DS, Taylor RE. Scientific Discovery at the Press of a Button: Navigating Emerging Cloud Laboratory Technology. Advanced Materials Technologies. 2024;9(16):2400084. https://doi.org/10.1002/admt.202400084

11. Adam L, McArthur GHt. Substitution Attacks: A Catalyst to Reframe the DNA Manufacturing Cyberbiosecurity Landscape in the Age of Benchtop Synthesizers. Appl Biosaf. 2024;29(3):172-80. https://doi.org/10.1089/apb.2023.0035

12. Wittmann BJ, Alexanian T, Bartling C, Beal J, Clore A, Diggans J, et al. Toward AI-Resilient Screening of Nucleic Acid Synthesis Orders: Process, Results, and Recommendations. bioRxiv. 2024:2024.12.02.626439.

13. Bari N, MacIntyre R. Smuggling biological materials and illegal laboratories–implications for biosecurity and potential biological attacks. Global Biosecurity. 2025. https://doi.org/10.31646/gbio.325

14. U.S. Department of Justice OoPA. Three Chinese National Scholars from University of Michigan Laboratory Charged for Conspiring to Smuggle Biological Materials into the U.S. 2025.

15. Majidi V. Statement of Vahid Majidi, Assistant Director, Weapons of Mass Destruction Directorate, Federal Bureau of Investigation, before the Committee on Homeland Security and Governmental Affairs, United States Senate, entitled "Ten Years After 9/11 and the Anthrax Attacks: Protecting Against Biological Threats". Testimony. Washington, DC: U.S. Department of Justice; 2011 2011/10/18.

16. Carter SR. Developing a Customer Screening Framework for the Life Sciences. Science Policy Consulting LLC; 2024 2024/03.

17. Matthews D. EU urged to consider DNA order screening to prevent an engineered pandemic. Science|Business. 2025 2025-01-16 01/16.

18. Carus WS. Bioterrorism and Biocrimes: The Illicit Use of Biological Agents Since 1900. Working paper. Washington, D.C.: Center for Counterproliferation Research, National Defense University; 2001 2001-02.

19. Washington MA, Paulino J, Kua S-C. Barriers to Biological Weapons Development: Potential Implications for Pathway Disruption. Countering WMD Journal. 2024(28):19-27.

20. Suk JE, Vogel KM, Ozin AJ. Dual-use life science research and biosecurity in the 21st Century: Social, Technical, Policy, and Ethical Challenges: Frontiers Media SA; 2015.

21. Kim HJ, Kim H-J. Yeast as an expression system for producing virus-like particles: what factors do we need to consider? Letters in Applied Microbiology. 2017;64(2):111-23. https://doi.org/10.1111/lam.12695

22. Wu Y, Gao S, Liu G, Wang M, Tan R, Huang B, Tan W. Development of viral infectious clones and their applications based on yeast and bacterial artificial chromosome platforms. Mol Biomed. 2025;6(1):26. https://doi.org/10.1186/s43556-025-00266-7

23. Thi Nhu Thao T, Labroussaa F, Ebert N, V'Kovski P, Stalder H, Portmann J, et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. Nature. 2020;582(7813):561-5. https://doi.org/10.1038/s41586-020-2294-9

24. Miest TS, Yaiw K-C, Frenzke M, Lampe J, Hudacek AW, Springfeld C, et al. Envelope-chimeric Entry-targeted Measles Virus Escapes Neutralization and Achieves Oncolysis. Molecular Therapy. 2011;19(10):1813-20. https://doi.org/10.1038/mt.2011.92

25. Louten J. Virus Structure and Classification. 2016.

26. Shao B, Yan J. A long-context language model for deciphering and generating bacteriophage genomes. Nature Communications. 2024;15(1):9392. https://doi.org/10.1038/s41467-024-53759-4

27. Twist B. A CRE.AI.TIVE application of AI: Engineering a More Resilient Global Food Supply. San Francisco, CA: Twist Bioscience; 2025 2025/06.

28. Hwang WY, Song J, Choe J, Ku KB, Kim H-S, Yoon GY, et al. De novo design of a safe and potent respiratory syncytial virus immuno-focusing antigen. bioRxiv. 2026:2026.01.28.702448.

29. GÃķtting J, Medeiros P, Sanders JG, Li N, Phan L, Elabd K, et al. Virology Capabilities Test (VCT): A Multimodal Virology Q&A Benchmark. arXiv preprint arXiv:250416137. 2025.

30. Polat S, Alyanak B, Dede BT, Temel MH, Yildizgören MT, Bağcier F. Marked reduction in hallucination rates with GPT-5: A positive development for medical and scientific writing. Malays Fam Physician. 2025;20:75. https://doi.org/10.51866/lte.1004

31. Zhang Y, Li M, Han WJ, Yao Y-F, Cen Z, Zhao D, editors. Safety is Not Only About Refusal: Reasoning-Enhanced Fine-tuning for Interpretable LLM Safety. Annual Meeting of the Association for Computational Linguistics; 2025.

32. Provatas K, Self A, Mouratidis I, Georgakopoulos-Soares I. BioVeil MATRIX: Uncovering and categorizing vulnerabilities of agentic biological AI scientists2026.

33. Park PS, Goldstein S, O'Gara A, Chen M, Hendrycks D. AI deception: A survey of examples, risks, and potential solutions. Patterns (N Y). 2024;5(5):100988. https://doi.org/10.1016/j.patter.2024.100988

34. Carlini N, Cheng N, Keane Lucas MM, Nasr M, Prabhushankar V, Xiao W, et al. Assessing claude mythos preview's cybersecurity capabilities. Anthropic Blog Red anthropic com. 2026.

35. Center for AIS, Innovation. Evaluation of DeepSeek AI Models. Gaithersburg, MD: National Institute of Standards and Technology; 2025 2025/09/30.

36. Wei B, Che Z, Li N, Sehwag U, Götting J, Nedungadi S, et al. Best Practices for Biorisk Evaluations on Open-Weight Bio-Foundation Models2025.

37. Sandbrink JB, Koblentz GD. Biosecurity risks associated with vaccine platform technologies. Vaccine. 2022;40(17):2514-23. https://doi.org/10.1016/j.vaccine.2021.02.023

38. Council NR, Sciences BoL, Sciences BoC, Policy BoHS, Systems CoEoNB, BioWatch, System tPH. BioWatch and public health surveillance: Evaluating systems for the early detection of biological threats: Abbreviated version: National Academies Press; 2011.

39. Engineering Biology Research Consortium Security Working G. Security Screening in Synthetic DNA Synthesis: Recommendations for Updated Federal Guidance. Engineering Biology Research Consortium; 2022.

40. Engineering Biology Research Consortium. Strengthening a Safe and Secure Nucleic Acid Synthesis Ecosystem: Outcomes of EBRC Stakeholder Engagement. Engineering Biology Research Consortium; 2025.

41. International Gene Synthesis C. IGSC Harmonized Screening Protocol v3.0. International Gene Synthesis Consortium; 2024 2024/09/03.

42. Hoffmann SA, Diggans J, Densmore D, Dai J, Knight T, Leproust E, et al. Safety by design: Biosafety and biosecurity in the age of synthetic genomics. iScience. 2023;26(3):106165. https://doi.org/10.1016/j.isci.2023.106165

43. Kane A, Parker MT. Screening State of Play: The Biosecurity Practices of Synthetic DNA Providers. Appl Biosaf. 2024;29(2):85-95. https://doi.org/10.1089/apb.2023.0027

44. Tarangelo JP, Attal-Juncqua A, Somani E, Roberts D, Webster K. Protecting Biological Materials and Services from Misuse: Opportunities for Access Monitoring and Control. Santa Monica, CA: RAND

Corporation; 2025 2025/10/20. Contract No.: RR-A4067-1.

45. The White House. Improving the Safety and Security of Biological Research Washington, DC: The White House; 2025 [updated 2025-05-05 2025/10/04. Available from: https://www.whitehouse.gov/presidential-actions/2025/05/improving-the-safety-and-security-of-biological-research/.

46. Diggans J, Leproust E. Next steps for access to safe, secure DNA synthesis. Frontiers in bioengineering and biotechnology. 2019;7:86. https://doi.org/10.3389/fbioe.2019.00086

47. Hassabis D, Altman S, Amodei D, Wang A, Graham P, Suleyman M, et al. In Support of Mandatory Nucleic Acid Synthesis Screening and Recordkeeping. An Open Letter. 2026.

48. Reason J. Human error: models and management. Bmj. 2000;320(7237):768-70. https://doi.org/10.1136/bmj.320.7237.768

49. Baum C, Berlips J, Chen W, Cui H, Damgard I, Dong J, et al. A system capable of verifiably and privately screening global dna synthesis. arXiv preprint arXiv:240314023. 2024.

50. Faure AJ, Martí-Aranda A, Hidalgo-Carcedo C, Beltran A, Schmiedel JM, Lehner B. The genetic architecture of protein stability. Nature. 2024;634(8035):995-1003. https://doi.org/10.1038/s41586-024-07966-0

51. Wimmer E, Mueller S, Tumpey TM, Taubenberger JK. Synthetic viruses: a new opportunity to understand and prevent viral disease. Nat Biotechnol. 2009;27(12):1163-72. https://doi.org/10.1038/nbt.1593

52. The White House. Winning the Race: America's AI Action Plan. Washington, DC: The White House; 2025 2025-07.

53. Williams A, Popescu S, Berke A, Vazquez E, Nevo S. Identifying and Closing Gaps in the Federal Select Agent Program: Opportunities for Improvement in an Era of Emerging Biotechnologies. Santa Monica, CA: RAND Corporation; 2025.

54. Mo W, Vaiana CA, Myers CJ. The need for adaptability in detection, characterization, and attribution of biosecurity threats. Nat Commun. 2024;15(1):10699. https://doi.org/10.1038/s41467-024-55436-y

55. Zhang Z, Zhou Z, Jin R, Cong L, Wang M. GeneBreaker: Jailbreak Attacks against DNA Language Models with Pathogenicity Guidance2025.

56. Gretton D, Wang B, Edison R, Foner L, Berlips J, Vogel T, et al. Exact-match search with functional variant prediction enables automated DNA screening. bioRxiv. 2025:2024.03.20.585782.

57. Bisardi M, Rodriguez-Rivas J, Zamponi F, Weigt M. Modeling Sequence-Space Exploration and Emergence of Epistatic Signals in Protein Evolution. Molecular Biology and Evolution. 2021;39(1). https://doi.org/10.1093/molbev/msab321

58. Bushuiev A, Bushuiev R, Sedlar J, Pluskal T, Damborsky J, Mazurenko S, Sivic J. Revealing data leakage in protein interaction benchmarks. arXiv preprint arXiv:240410457. 2024.

59. Chopra G, Summa CM, Levitt M. Solvent dramatically affects protein structure refinement. Proc Natl Acad Sci U S A. 2008;105(51):20239-44. https://doi.org/10.1073/pnas.0810818105

60. Kilinc M, Jia K, Jernigan RL. Improved global protein homolog detection with major gains in function identification. Proceedings of the National Academy of Sciences. 2023;120(9):e2211823120. https://doi.org/10.1073/pnas.2211823120

61. Nerín-Fonz F, Cournia Z. Machine learning approaches in predicting allosteric sites. Current Opinion in Structural Biology. 2024;85:102774. https://doi.org/10.1016/j.sbi.2024.102774

62. Olanders G, Testa G, Tibo A, Nittinger E, Tyrchan C. Challenge for Deep Learning: Protein Structure Prediction of Ligand-Induced Conformational Changes at Allosteric and Orthosteric Sites. J Chem Inf Model. 2024;64(22):8481-94. https://doi.org/10.1021/acs.jcim.4c01475

63. Zhao Q, Zhao H, Guo L, Zheng K, Li Y, Ling Q, et al. ColdstartCPI: Induced-fit theory-guided DTI predictive model with improved generalization performance. Nat Commun. 2025;16(1):6436. https://doi.org/10.1038/s41467-025-61745-7

64. Robertson WE, Rehm FBH, Spinck M, Schumann RL, Tian R, Liu W, et al. Escherichia coli with a 57-codon genetic code. bioRxiv. 2025:2025.05.02.651837.

65. Flamholz ZN, Biller SJ, Kelly L. Large language models improve annotation of viral proteins. Res Sq. 2023.

66. Lauber C, Seitz S. Opportunities and Challenges of Data-Driven Virus Discovery. Biomolecules. 2022;12(8). https://doi.org/10.3390/biom12081073

67. Yassif J, Severance H, Isaac C. Tabletop Exercise Summary: Calling for Bold Action to Prevent the Next Biological Catastrophe. Washington, DC: Nuclear Threat Initiative (NTI | bio); 2023 2023/05/25.

How to cite this article: Thottunkal S.Artificial Intelligence Enabled Viral Weapons of Mass Destruction, A Critical Perspective on The Current Landscape of Nucleic Screening, AI Agents, and Implementing the National Artificial Intelligence Action Plan. Global Biosecurity. 2026; 8(1).

Published: June 2026