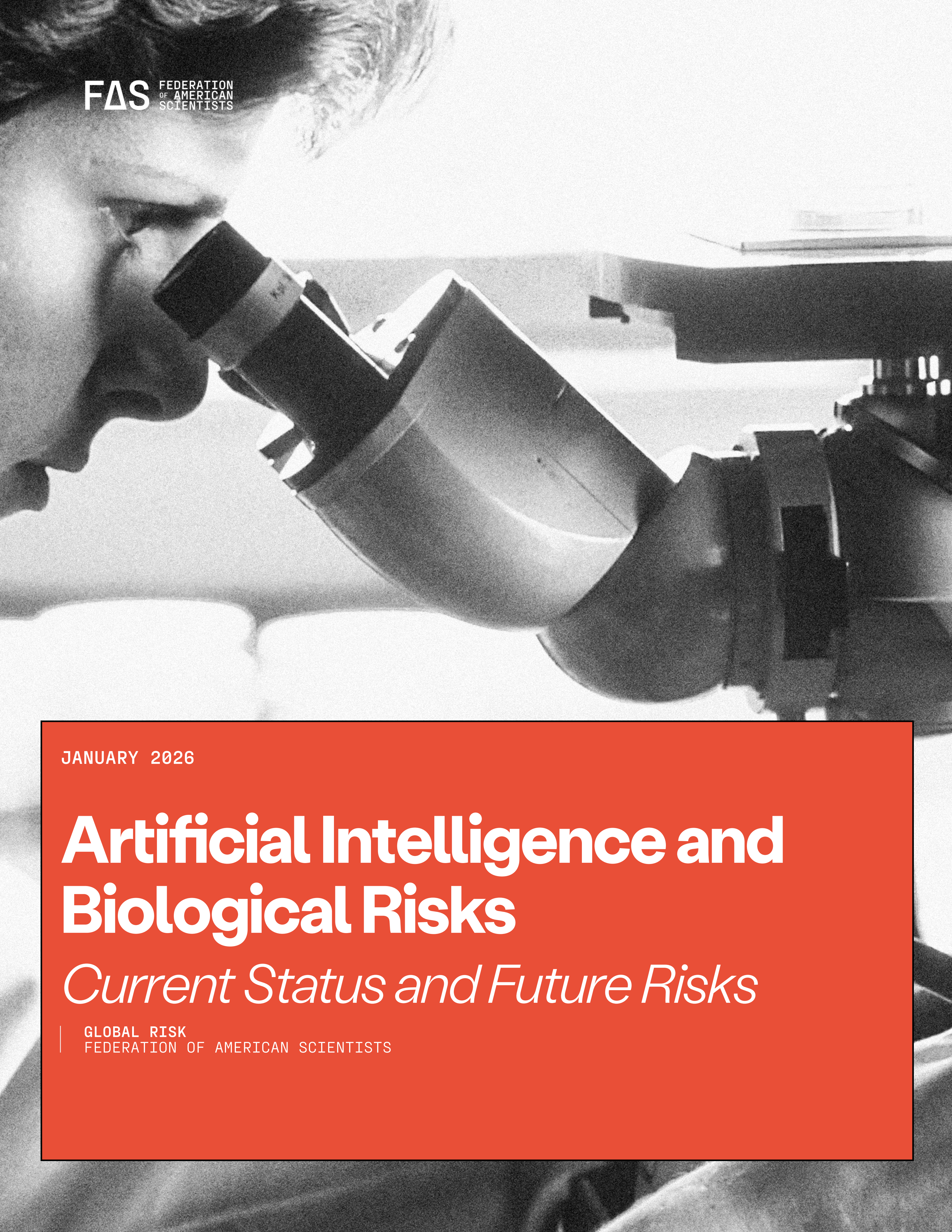

FAS FEDERATION OF AMERICAN SCIENTISTS

JANUARY 2026

Artificial Intelligence and Biological Risks

Current Status and Future Risks

GLOBAL RISK
FEDERATION OF AMERICAN SCIENTISTS

ABOUT THIS REPORT

This report summarizes the key findings, insights, and policy options from a June 2025 roundtable event on risks at the convergence of artificial intelligence (AI) and biology. The Federation of American Scientists, in partnership with the Future of Life Institute (FLI), brought together academic, industry, and government experts spanning AI, biology, and technology policy domains for this conversation. Experts identified bottlenecks in how we understand and identify threats at the AIxBio intersection (particularly deliberate misuse), took stock of advancing AI capabilities and their impact on bioweapons and accidental releases, and considered opportunities to avert risk in even a geopolitically heating world.

This report is structured in three parts: an executive summary, a detailed analysis of the findings, and three papers authored for participants in advance of the event. Hamza Chaudhry, AI and National Security Lead at FLI, authored the first pre-read on conceptualizing and framing risks in biology, AI, and the AIxBio convergence. Dr. Oliver Stephenson, Associate Director of AI and Emerging Technology Policy at FAS, authored the second pre-read on risk mitigation approaches in biology, AI, and the AIxBio convergence. Dr. Yong-Bee Lim, Associate Director of Global Risk at FAS, authored the final pre-read on how experts frame future risk in the bio, AI, and AIxBio domains, and the opportunities some of this framing provides us for greater safety and security.

Global Risk Program at FAS

The Global Risk Program at the Federation of American Scientists (FAS) focuses on addressing and preventing the events and threats that could permanently cripple or destroy humanity. Some key areas our team focuses on include nuclear war, the next global pandemic, biological attack, and even a collision with a massive near-earth object. Our team of policy experts, scientists, and researchers use tools including forecasting, research and analysis, and expertise in key global risk domain areas to develop modern policy solutions for a rapidly advancing and complex time in humanity's development. Find out more at our website www.fas.org/issue/global-risk. The project is led by Jon B. Wolfsthal the Director of the Global Risk Program at FAS.

Funding

This report and the associated workshop were made possible through the generous support of the Future of Life Institute and are part of a wider series in our ongoing "AIxGlobal Risk Nexus" project. This project will culminate in a global summit in Spring 2026. The views expressed in this report are those of the authors and do not necessarily reflect the positions of the funders or participants.

Special thanks to Yong-Bee Lim, PhD, Associate Director of the Global Risk Program, Oliver Stephenson, PhD, Associate Director of Artificial Intelligence and Emerging Technology Policy, Elliott Gunnell, M.Sc. Project Associate, Global Risk Program, and Abhay Katoch, Visiting Fellow with the Global Risk Program, for their contribution to the event and this report. Special thanks to our colleague on FAS's Communications team, Kate Kohn, for developing the graphic for this report.

FAS can be reached at 1150 18th St. NW. Suite 1000, Washington, DC, 20036, **fas@fas.org**, or through **fas.org.**

CONTENTS

EXECUTIVE SUMMARY

On June 12, 2025, the Federation of American Scientists (FAS), in partnership with the Future of Life Institute (FLI), convened a D.C.-based roundtable at the Lyle Hotel on risks emerging at the convergence of AI and bio. Discussions examined how AI and biosecurity communities define and assess risk, where rapid advances in both domains may create new pathways for deliberate misuse or accidental harm, and which risk-reduction opportunities remain feasible amid heightened geopolitical competition, including between the United States and China. Participants included technical and policy experts from academic, industry, and government, including representatives from the Offices of Senator Chuck Schumer (D-NY) and Representative Chrissy Houlahan (D-PA-6), as well as officials from the National Security Commission on Emerging Biotechnology (NSCEB).

FINDINGS

Participants broadly agreed that advances in biology and AI are reshaping the biological risk landscape. Several emphasized that general-purpose systems—particularly large language models (LLMs)—can lower bio-relevant knowledge and problem-solving for a wider range of users, including potentially malicious actors. Others focused on specialized AI-enabled biological research tools, including biological design tools (BDTs), which may raise the ceiling on what sophisticated actors can do. Together, these dual dynamics—risk from diffusion and risk from amplification—complicate the "one-size-fits-all" approaches to governance we historically see in the weapons of mass destruction (WMD) domain.

Participants repeatedly returned to a core governance bottleneck: the speed and structure of current testing and evaluation for frontier models lag the pace of capability development. This gap is creating a growing mismatch between what is known—or can be demonstrated—about model performance and what frontier systems can do in practice. Experts noted that this challenge is compounded by biology's measurement constraints. Uncertainty in biological data, variability across geographic and environmental contexts, and the role of tacit knowledge in laboratory work all complicate efforts to rigorously quantify AI-driven improvements in biological capability, or "AI uplift". Taken together, these evaluations and data limitations make it harder to set credible baselines, track change over time, and design mitigations that remain valid as capabilities evolve.

Despite uncertainty in timelines and pathways, many participants expressed confidence that meaningful risk reduction is achievable today. Experts converged on the value of **layered defenses** spanning both the AI and bio ecosystems, including: 1) AI model safeguards and access controls; 2) synthesis screening and other points of contact in biological supply chains; 3) monitoring and incident response; and 4) preparedness efforts that reduces consequence even when prevention fails.

Participants also flagged a perennial structural challenge: responsibilities are fragmented across agencies, jurisdictions, and the private sector, while much of the relevant innovation occurs outside government. This fragmentation increases the importance of clearer expectations, practical guidance, and sustained channels for technical collaboration and information-sharing through public-private partnerships.

Finally, participants underscored that the international context matters: risk reduction should be pursued through approaches that remain viable amid geopolitical tension, including pragmatic engagement, norm-setting, and scientific venues where appropriate.

POLICY OPTIONS

Based on roundtable discussion and subsequent analysis, FAS identified the following policy options to address challenges at the nexus of AI and biology. These options are included for discussion purposes, and their inclusion does not imply endorsement by any participants of the roundtable discussions:

- **Create a rapid AI×Bio evaluation capability** to measure AI-enabled “uplift” in biologically relevant tasks and track capability change over time.
- **Resource a sustained research agenda on AI×Bio risk**, including benchmarks and real-world studies, to address the evidence gap constraining governance.
- **Strengthen and scale DNA synthesis screening** as a durable defense at the digital-physical boundary, including improved customer and order risk analysis.
- **Adopt risk-tiered managed access for bio-relevant AI capabilities**, pairing safeguards with monitoring and anomaly detection.
- **Reduce fragmentation in federal AI×Bio governance** by improving government coordination and clarifying expectations for industry.

WHAT WE HEARD

On June 12th, 2025, the Federation of American Scientists (FAS), in partnership with the Future of Life Institute (FLI), convened a Washington, D.C. roundtable bringing together over 50 experts from academia, industry, and government to examine risks at the convergence of AI and biology. The purpose of the roundtable was to explore how AI and biosecurity communities conceptualize and assess risk, how rapidly advancing capabilities in both domains may create new pathways for deliberate misuse or accidental harm, and which risk-reduction opportunities are available even amid heightened geopolitical competition.

This discussion occurred during a period of accelerating change in both AI development and the life sciences. AI systems are already being deployed at scale in high-consequence, public-sector environments—illustrating both rapid diffusion and the challenges of keeping policy, evaluation, and oversight aligned with real-world use. For example, the Department of Defense's Chemical and Biological Defense Program (CBDP) and the Defense Innovation Unit (DIU) publicly announced an initiative to strengthen AI-driven biosurveillance capabilities, reflecting growing interest in applying advanced analytics and AI towards detecting and characterizing biological threats further "left of boom."[1] At the same time, with AI tools increasingly integrated into life science workflows, policymakers will need governance approaches that remain credible as capabilities diffuse and change rapidly.

FINDING 1. AI IS RESHAPING THE BIOLOGICAL RISK LANDSCAPE BY LOWERING BARRIERS AND RAISING THE CEILING

Participants emphasized that advances in increasingly general-purpose AI are reshaping the biological risk landscape in ways that strain existing biosecurity and biosafety governance assumptions. Several noted that many risk frameworks implicitly assume that high-consequence biological capabilities are concentrated among a relatively small set of well-resourced actors and institutions. By contrast, general-purpose models—especially widely accessible LLMs—may be lowering barriers to bio-relevant knowledge, planning, and troubleshooting, potentially expanding the pool of users able to engage with sensitive biological activities. While this does not automatically translate into real-world harm, it does complicate how experts and policymakers work together to think about prevention, detection, and consequence management at the AIxBio nexus.[2]

Participants also highlighted that the AIxBio convergence is not solely a diffusion problem; it may also raise the ceiling of what sophisticated actors can design or attempt. Discussion pointed to the emergence of biological design tools (BDTs) that can accelerate elements of scientific discovery and engineering by assisting with design choices, prioritizing experiments and workflow, or rapidly troubleshooting errors while implementing a protocol.[3] This acceleration, some noted, could compress iteration cycles and reduce the time and expertise needed to advance along potentially harmful research pathways. Others cautioned that biology remains constrained by tacit knowledge, experimental variability, and data limitations. Therefore, it is difficult to make confident, generalizable claims about how quickly AI will translate into "uplift" for risk-relevant tasks.

The discussion underscored that these diffusion and amplification dynamics increase the importance of evaluating AIxBio risk at the level of real-world systems rather than any single model or platform. Participants noted that bio-relevant capabilities are also increasingly connected to other systems, complicating attempts to rely on

[1] U.S. Department of Defense's Defense Innovation Unit, "CBD and DIU Strengthen National Security with AI-Driven Biosurveillance Initiative," Defense Innovation Unit, accessed January 15, 2026, https://www.diu.mil/latest/cbd-and-diu-strengthen-national-security-with-ai-driven-biosurveillance

[2] Christopher A. Mouton, Caleb Lucase, and Ella Guest, The Operational Risks of AI in Large-Scale Biological Attacks: Results of a Red-Team Study, Research Report RRA2977-2 (Santa Monica, CA: RAND Corporation, 2024).

[3] National Security Commission on Emerging Biotechnology, White Paper 3: Risks of AIxBio (Washington, D.C.: National Security Commission on Emerging Biotechnology, January 2024), https://www.biotech.senate.gov/wp-content/uploads/2024/01/AIxBio-White-Paper-3-1.pdf

narrow, model-specific mitigations. These dynamics highlight that as AI tools become more embedded in life-science research and development, government will need to be calibrated across a wider range of actor types and contexts, including settings outside of federal funding streams and traditional public health, health security, or biodefense authorities.

At the same time, participants emphasized that the risks posed by diffusion and amplification are not insurmountable, but require a shift toward layered defenses that can perform under uncertainty. Suggested approaches included strengthening baseline safeguards for broadly deployed models (e.g., tighter refusal behavior for illicit biosecurity requests), implementing more rigorous oversight mechanisms for higher-risk BDTs, and investing in model-adjacent measures that reduce consequences even when prevention fails (e.g., improved biosurveillance, incident response capacity, and preparedness). In short, participants converged on the concept that because AI may both broaden access and amplify capability, the policy objective should be to build a resilient, multi-layered risk management posture: one that is technically credible, operationally implementable, and durable even as frontier capabilities evolve.

FINDING 2. THE BINDING CONSTRAINT FOR RISK REDUCTION IS EVIDENCE: TESTING AND EVALUATION ARE NOT KEEPING PACE WITH FRONTIER CAPABILITY

Participants emphasized that advances in frontier AI are outpacing the current evidence base for understanding how these systems affect biological risk in the real world. Several attendees noted that "testing and evaluation" for AIxBio-relevant capabilities—especially claims about user "uplift"—often lags behind the tempo of model development and development. Indeed, one participant highlighted that AIxBio research shaping policy understanding in 2025 was published in 2024, using AI models created in 2022. This three-year period has seen incredibly rapid AI progress.

The relative pace of AI progress compared to AI evaluations creates a widening mismatch between what is publicly known about model performance and what real-world systems may be able to do. Therefore, experts and decision-makers lack a stable, replicable method for answering policy-relevant questions about capability change over time: what new capabilities are emerging, who can access them, under what conditions they translate into physical effects, and which mitigations remain robust (or fail) as models and biological workflows evolve. This concern mirrors broader federal guidance emphasizing that trustworthy AI requires continuous measurement and management across the lifecycle.[4]

Participants also highlighted how biology-specific constraints compound these evaluation challenges. Even where a model can produce plausible technical outputs, translating those outputs into reliable, replicable biological outcomes is shaped by uneven data quality, context-specific variability, and tacit laboratory and operational knowledge.[5] This introduces a gap that several attendees stressed makes it difficult to produce clean, generalizable "uplift" estimates—especially when uplift depends on access to equipment, materials, and tacit knowledge rather than information alone. Therefore, policymakers would be wise to avoid overconfidence in either direction: assuming that impressive-looking outputs necessarily translate into real-world capability, or assuming that today's bottlenecks will reliably hold in the future.

Several attendees also noted that evaluation is increasingly challenged by how bio-relevant capability often emerges at the level of systems. Participants pointed to tool-chaining (front-to-end automated workflows), workflow integration, and the ability for users to browse across different AI models further muddy our

[4] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0) (Gaithersburg, MD: NIST, 2023)

[5] Filippa Lentzos, Jez Littlewood, Hailey Wingo, and Alberto Muti, "Apathy and Hyperbole Cloud the Real Risks of AI Bioweapons," Bulletin of the Atomic Scientists, September 12, 2024, https://thebulletin.org/2024/09/apathy-and-hyperbole-cloud-the-real-risks-of-ai-bioweapons/

understanding of real-world risk. Further, in practice, participants raised that evaluation approaches should become more operationally realistic: they should measure capabilities in contexts that approximate how models are actually used, clarify assumptions about user access and limitations, and be paced such that it tracks frontier model iteration.

At the same time, participants emphasized that the evidence gaps that currently exist in assessing AIxBio risks can be addressed if policymakers treat evaluation as a core component, rather than as an afterthought. Several participants suggested that a credible "uplift" evidence base will require: 1) sustained investment in measurement methods; 2) structured collaboration among AI developers, biosecurity experts, and other stakeholders; and 3) bifurcated pathways for sharing results, where public findings are provided for transparency and democratic accountability, and protected channels are used to share sensitive details with select individuals and communities.

FINDING 3. RISK REDUCTION IS FEASIBLE NOW, BUT IT REQUIRES LAYERED DEFENSES AND STRONGER CONNECTIVE TISSUE

Participants raise that, despite uncertainty about timelines and specific misuse pathways, meaningful risk reduction at the AIxBio nexus is achievable today. Rather than treating AIxBio as a problem that can be solved through a single safeguard, several attendees gravitated towards a multi-layered, or "defense-in-depth" approach across both AI and biological ecosystems. Participants noted that model-level interventions, such as safeguards and access controls, are necessary **but insufficient** to prevent malicious actors from leveraging different tools and access regimes from potentially achieving their ends. This reflects a basic governance reality: because AI may lower barriers for a broader set of users while simultaneously raising the ceiling for sophisticated actors, the most durable posture is one that reduces risk at multiple points:

- **Upstream:** Examples include model behavior and limiting access to information or capabilities.
- **Midstream:** Examples include enabling services and supply chains, such as gene synthesis and sequencing companies.
- **Downstream:** Examples include detection, response, and consequence management systems.

Participants highlighted that, currently, the most actionable layers are "model-adjacent", where prevention measures taken outside of the model may contribute significantly to risk reduction. Several pointed to domains like synthesis screening and supply-chain friction points as particularly meaningful, as they can raise the cost of misuse, increase the probability of detection, and create clear pathways for escalation and reporting. This aligns with recent U.S. policy moves that seek to formalize and scale nucleic acid synthesis screening, including the White House OSTP framework and NIH funding-linked requirements related to synthetic nucleic acid procurement.[6] These initiatives provide a "carrots-and-sticks" model that incentivizes broad and multiple targets of opportunities for risk reduction activities.

Participants also stressed that monitoring and incident response are underdeveloped relative to the pace of technological change, and that preparedness is a critical layer because prevention can fail. Multiple participants emphasized that risk reduction strategies should not depend solely on upstream controls, but should include mechanisms to detect anomalies early, including potentially connected anomalies across the AIxBio ecosystems: strategies that some referred to as "no-regret" capacity since they reduce consequence across a range of threat scenarios.

[6] For sources, see White House Office of Science and Technology Policy, Framework for Nuclear Acid Synthesis Screening (Washington, D.C.: OSTP, April 29, 2024); and National Institutes of Health, "Notification of NIH Requirements for the Safe, Secure, and Responsible Procurement of Synthetic Nucleic Acids," NOT-OD-25-023 (October 25, 2024).

However, existing preparedness efforts have limitations. For example, the U.S. government's BioWatch program—designed to detect the early release of biological agents in cities—has struggled with false positives, often triggered by naturally occurring microbes rather than man-made weapons. These false positives undermined trust in the system and complicated rapid response, illustrating how detection and preparedness tools can fail if they are not calibrated to real-world biological variability.[7]

At the same time, participants emphasized that layered defenses will not scale without stronger connective tissues across government, industry, and the scientific community. Several attendees identified fragmented responsibility and a lack of AIxBio champions as part of what has held back biological risk reduction efforts. As a result, participants highlighted the importance of clearer expectations, practical guidance that is actually implementable, and sustained channels for technical collaboration and information-sharing with industry.

Finally, participants underscored that the international context matters: risk reduction measures remain viable amid geopolitical tension and uneven adoption across countries and industries around the world. Several noted that pragmatic norm-setting through scientific venues and Track 1.5 or Track 2.0 expert engagement can still advance safety and security, even when formal international mechanisms move slowly.

[7] U.S. Government Accountability Office, DHS Exploring New Methods to Replace BioWatch and Could Benefit from Additional Guidance (Washington, D.C.: U.S. Government Accountability Office, May 2021), https://www.gao.gov/assets/720/714434.pdf

MENU OF POLICY OPTIONS

Based on the panel discussions and subsequent analysis, we extracted the following menu of options which policymakers could use to address challenges at the nexus of AI and Biology. These options are included for discussion purposes, and their inclusion does not imply endorsement by any participants of the roundtable discussions:

The U.S. government could **establish an ongoing agile AIxBio evaluation and stress-testing program**—modeled on rapid, iterative testing—focused on measuring uplift. This effort should be paired with incorporating the right biological expertise to bridge gaps that exist on both the biological and AI community sides.

Congress could **resource a sustained, evidence-generating "uplift" research agenda** that reduces today's measurement bottleneck. This would enable funding lower-cost benchmarks and real-world studies and increase confidence in AIxBio risk assessments as frontier capabilities evolve. Such efforts would also bolster ongoing U.S. government evaluation efforts suggested in the policy option above.

The U.S. government could **modernize and strengthen DNA synthesis screening** as a core component of layered defenses at the AIxBio digital-physical divide. This includes a broader uptake of screening best practices, customer and order risk analysis, and clearer communication and escalation pathways.

Model developers could adopt **risk-tiered** "managed access" and monitoring systems for bio-relevant capabilities, which would pair safeguards and refusals with graduated access controls, logging, and anomaly detection.

The Executive Branch of the U.S. government **could reduce fragmentation** by consolidating and coordinating AIxBio risk responsibilities and guidance.

CONCLUSION

The roundtable underscored that AI is changing the biological risk landscape in ways that are rapid and difficult to bound within legacy biosecurity frameworks. However, many of the most consequential risk pathways remain speculative. This uncertainty makes it difficult to confidently measure the degree of "uplift" that frontier models and biological design tools may provide to users.

Despite this disagreement, participants emphasized that if risks do concretely emerge, they are likely to do so through two simultaneous pathways: **diffusion** of bio-relevant knowledge and capability to broader populations, and **amplification** of what sophisticated actors can design or attempt. In practice, this means policymakers should resist a one-size-fits-all approach imported from traditional WMD governance and instead adopt approaches calibrated to distinct actors, contexts, and points of leverage across the AIxBio ecosystem.

At the same time, the discussions made clear that the binding constraint for governance is emerging: testing and evaluation of frontier systems are not keeping pace with model improvements. Biology's data limitations, variability across contexts, and reliance on tacit knowledge competencies complicate efforts to quantify the "uplift" AI tools and systems lend to life sciences activities with confidence.

Importantly, participants expressed confidence that meaningful risk reduction is achievable now if done in a "defense-in-depth" fashion. Model safeguards and managed access are necessary, but they can be supported by model-adjacent activities that remain effective even when model-level controls may fail. This posture is not only technically realistic; it is also implementable under uncertainty, and can even accommodate dynamics like uneven adoption and rapid capability change.

Finally, participants emphasized that durable progress will depend on building stronger connective tissues across government, industry, and the scientific community. This can even be done amid geopolitical competition through the establishment of durable practices, repeatable evaluation mechanisms, and layered safeguards that reduce risk across a wide range of futures while preserving the benefits and opportunities of responsible innovation.

PRE-READ PAPERS FROM ROUNDTABLE

CONCEPTUALIZING AND FRAMING RISKS IN BIO, AI, AND THE BIO×AI CONVERGENCE

MR. HAMZA CHAUDHRY, FUTURE OF LIFE INSTITUTE

Biotechnology and artificial intelligence are converging in ways that could transform both public health and national security. This pre-read outlines how biological risks have evolved and how AI may reshape the biosecurity landscape for policymakers. It begins with a brief history of biosecurity concerns, then reviews key modern biological risks (e.g., synthetic biology, dual-use research, pathogen access) and how they have changed over the past few years. It then examines how AI — from large language models to protein prediction tools — is altering the risk landscape, both improving capabilities and introducing new threats. Finally, it highlights emerging concerns at the AI-biosecurity nexus, such as AI-assisted design of dangerous pathogens, dual-use datasets, and difficulties in risk assessment and control, and concludes with questions for discussion.

Historical Context

Innovations in biotechnology have long been a national security concern. In the 20th century, major powers pursued biological weapons programs, prompting international bans like the 1972 Biological Weapons Convention. The Soviet Union's massive Cold War bioweapons program shocked the world, and although such large-scale state programs waned, worries persisted. The 2001 anthrax mail attacks and terrorist attempts (e.g., the 1984 Rajneeshee cult poisoning) underscored the threat of bioterrorism by non-state actors. Fortunately, most non-state attempts failed due to the complexity of acquiring, growing, scaling, and weaponizing pathogens. Policymakers also grew alert to "dual-use" research – legitimate science that could be misused – after controversial experiments (such as the 2011 H5N1 flu gain-of-function studies) showed how scientific advances might inadvertently enable more lethal pathogens.

In the past decade, the number of high-containment laboratories (e.g., BSL-3/4 labs) worldwide has increased, and with it the frequency of dangerous incidents – possibly including the COVID-19 pandemic itself if a lab accident was involved. COVID-19 drove home the catastrophic potential of pandemics, causing at least 9 million deaths globally, exposing gaps in national preparedness and bringing biosecurity to the forefront as a policy issue about not just hostile actors but also lab safety, pandemic preparedness, and governance of emerging biotechnologies.

Evolving Risks in Biology and Biotechnology

Advances in the life sciences have opened unprecedented opportunities – and new risks – in recent years. Key areas of concern include:

Synthetic Biology & Genome Editing

The ability to engineer organisms has grown markedly. Techniques like CRISPR gene editing allow scientists to modify pathogens or create novel organisms with ease, unthinkable a decade ago. Synthetic biology efforts have already synthesized viruses from scratch – for example, researchers recreated horsepox virus (a relative of smallpox) using mail-order DNA for an estimated cost of about $100,000 in 2017. As DNA synthesis gets cheaper and more widespread, the barrier to assembling dangerous viruses is lowering. Such capabilities can yield positive breakthroughs (vaccines, engineered probiotics) but also enable an actor to resurrect eradicated viruses or construct novel pathogens.

Dual-Use Research and Gain-of-Function

Legitimate biomedical research can inadvertently create security risks. For instance, "gain-of-function" studies that enhance a pathogen (to study its transmissibility or virulence) have sparked debate. In the last five years, U.S. oversight of potentially pandemic pathogens has been reconsidered and strengthened in response to these concerns. Yet experiments demonstrating the ease of constructing or enhancing viruses continue to raise alarms. The line between beneficial research and dangerous knowledge is thin, requiring careful biosecurity and bioethics oversight. The National Security Commission on Emerging Biotechnology (NSCEB) in 2025 urged working with international partners to develop biosafety and biosecurity standards to prevent misuse, whether deliberate or accidental.

Pathogen Access & Pandemic Potential

The proliferation of high-end labs and global exchange of biological materials means more actors may be able to access dangerous pathogens. While top-tier bioweapons programs remain complex (states still find large-scale bioweapons of limited military utility due to controllability issues), the global diffusion of biotechnology know-how means a lone scientist or small group may more easily obtain or grow deadly agents. Black-market or illicit procurement of pathogens is a risk, but so is simply ordering DNA and using gene synthesis services if oversight fails. The existing framework of bioweapons conventions and lab safety norms leaves some gaps that could be exploited. Moreover, rapid advances in gene sequencing and data sharing mean that the genomic sequences of lethal pathogens (like 1918 influenza or Ebola) are publicly available – a resource for science, but also a potential "recipe" for misuse if coupled with synthesis technology.

Recent Trends

In the past five years, biotechnology has accelerated. The cost of DNA sequencing and synthesis has plummeted, global bio-innovation investment has surged, and biotechnology has been democratized by broader access to tools. More countries and even DIY community labs are involved in biotechnology. Positive developments (like mRNA vaccine platforms, gene therapies, and bio-based manufacturing) come alongside heightened risk of accidents or abuse. The U.S. government recognizes biotechnology as a strategic domain – for example, NSCEB's 2025 report calls for prioritizing biotech at the national level and treating biological data and infrastructure as strategic resources. In short, biological risk today is a moving target: it encompasses not just old specters of state bioweapons, but also new possibilities of small-scale but devastating attacks or accidents enabled by cutting-edge science.

How AI Is Changing the Biological Risk Landscape

Artificial intelligence is amplifying both the capabilities and risks in the biotech arena. Several AI advancements are particularly salient:

Large Language Models (LLMs) and Knowledge Access

The rise of AI systems like GPT-4 demonstrates that vast scientific knowledge can democratize access to scientific knowledge. These AI chatbots can potentially explain or even devise experimental methods in biology. There is concern that a non-expert with malicious intent might use an LLM to obtain step-by-step instructions for creating a pathogen or toxin. However, recent studies suggest that current AI models have limitations. A RAND-sponsored red-team exercise in 2023-24 found that today's LLMs did not significantly improve the viability of biological attack plans for non-state actors compared to using the internet alone. Indeed, in that experiment, experts saw no statistically significant difference in the quality of attack plans with or without LLM-based help. As AI grows more capable, it could close these knowledge gaps. RAND analysts warned that AI is advancing faster than governments can react, and even if GPT-4 didn't enable a bio-attack, future models might pose such a risk. OpenAI's own evaluations similarly found GPT-4 gave at most a "mild uplift" to users trying to formulate biological threats (a modest increase in accuracy/completeness that was not statistically significant). Twenty months later, OpenAI scored its O3 model as being 'medium' for CBRN risk, a higher bar for risk. In other words, AI assistance in this realm is currently marginal – but this situation is changing, and it signals a need to monitor model improvements closely.

AI in Protein Folding and Bioinformatics

AI has revolutionized our ability to predict and design biological molecules. DeepMind's AlphaFold AI, for example, has predicted the 3D structures of over 200 million proteins and released them in an open database. This is a triumph for science – accelerating drug discovery and bioengineering – but it also exemplifies the dual-use dilemma. By demystifying protein structures (including those of pathogens and toxins), AI tools can help design countermeasures (like new vaccines or antivirals) and potentially help optimize harmful biological agents. Researchers can more readily identify which parts of a virus are critical for function or immune evasion, information a bad actor could use to design a more potent variant. Moreover, AI-driven bioinformatics can sift through genomic data to find dangerous gene sequences or properties. For instance, algorithms might identify genetic markers that make a pathogen more deadly or help tailor a biological agent to target specific populations – raising the specter of AI-informed bioweapons targeting particular ethnic or genetic groups (though this remains technically very challenging with the current state of the art).

Generative AI for Molecular Design

AI systems are now used to design new molecules, from pharmaceuticals to enzymes. In benevolent hands, they can propose lifesaving drugs; in malevolent hands, they could conjure novel toxins. In 2022, researchers were able to easily repurpose a drug-discovery AI to generate dangerous chemical agents. Analogous capabilities in biology (e.g., designing proteins or DNA sequences) are emerging. AI could conceivably propose new virus mutations or synthetic genomes optimized for virulence. This autonomous creativity is powerful: it lowers the expertise and time needed to find deadly designs. As the RAND report noted, as AI advances it will "likely lower the barrier for all actors (including malign actors) … to conceptualize, plan, and conduct" chemical or biological attacks. In short, AI's ability to generate or optimize designs could supercharge the ideation phase of biothreat development.

Automation and "Autonomous Discovery" Platforms

Beyond analysis and design, AI is driving automation in laboratories. Robotic labs and cloud laboratory services allow experiments to be designed by software and executed remotely. AI planning tools can iterate over biological experiments rapidly. This has legitimate uses (speeding R&D, democratizing lab access), but it also means a would-be bad actor might not need hands-on expertise – they could rent a cloud lab, feed in AI-generated protocols, and produce a harmful agent from afar. Some cloud lab companies screen for illicit or dangerous experiment requests, but not all do. The combination of AI-driven planning and automated execution could remove traditional bottlenecks (like requiring advanced skills or physical lab access) that have historically protected us from amateur bioterrorists. For example, a future AI system might autonomously design an experiment to synthesize a pathogen, then send instructions to a contract lab robot. Even in more ordinary settings, AI can troubleshoot scientific protocols – identifying why an experiment failed and suggesting fixes – thereby accelerating the "design-build-test" cycle for developing a pathogen. This could shorten the time required to go from a concept to a viable bioweapon, should an adversary attempt it.

AI is a force-multiplier in biotechnology. It augments capabilities across the board: scientific discovery is faster (e.g., vaccines for COVID-19 were developed in record time, partly thanks to computational tools), but so is potential misuse. AI models are inherently dual-use: the same model that finds a cure can be turned to find a toxin. This convergence of AI and bio raises novel challenges for security policymakers, who must balance promoting innovation with guarding against catastrophe.

Emerging Concerns at the AI–Biosecurity Frontier

Several frontier issues demand attention as we consider policy responses at the Bio×AI convergence:

Malicious Use of AI Models ("AI as a Biothreat Enabler")

How far off is a scenario where a rogue actor asks an AI assistant and succeeds? We've seen hints: highly capable language models could theoretically provide step-by-step protocols, troubleshooting advice, or even automated execution for creating biological weapons. Tests indicate current models are not a silver bullet for would-be bioterrorists – much of the dangerous knowledge is already available in literature and on the internet. However, as models incorporate more specialized scientific data and reasoning ability, their usefulness to bad actors could increase. There is also the risk of fine-tuned models (for example, an open-source AI additionally trained on pathogen genomics or virology texts) that might be more willing and able to provide harmful guidance if safeguards fail. AI researchers have proposed developing tripwire evaluations – early warning tests to signal when an AI's capability to aid in biothreat creation becomes dangerously effective. The fundamental challenge is that AI tech is widely accessible and rapidly progressing, so any malicious use case discovered will be hard to contain globally.

Dual-Use Datasets and Information Hazards

AI development thrives on open data and shared knowledge, but in biosecurity, this collides with the concept of "information hazards" – certain knowledge that can itself be dangerous. For example, large public datasets of pathogen genomes, protein structures, or biochemical compounds are invaluable for research. Yet these same datasets can be used to train AI models that might generate harmful outputs. An AI model trained on the full genome sequences of eradicated or extinct pathogens (like smallpox or deadly flu strains) could potentially be prompted to "invent" a viable virus that evades current vaccines. Even publishing scientific findings poses dilemmas: a paper that openly details how to enhance a virus or the precise recipe for a toxin provides a blueprint that an AI could easily memorize and regurgitate. Despite innovations in AI model content filtering, fine-tuned refusals, or even training approaches that make models "forget" or avoid certain hazardous patterns, deciding what constitutes inherently dangerous knowledge and controlling its diffusion is a policy quandary. The international nature of science and AI means that unilateral restrictions may only have a limited effect if adversaries can access the same data or models.

Risk Assessment and Model Evaluation

Unlike traditional weapons, the weaponization potential of an AI algorithm is intangible and hard to quantify. How do we assess if a given AI is a biosecurity threat? The RAND red-team approach and OpenAI's controlled evaluations are some of the first attempts to measure how much AI could boost a malicious actor. These methods involve subject-matter experts simulating adversaries and seeing what difference the AI makes in planning an attack. So far, results show only marginal changes, but importantly, they establish baseline metrics for risk. Going forward, the government and industry will need to routinely red-team AI models for misuse scenarios, especially as new versions are more capable. One challenge is defining the threshold at which AI assistance becomes significant in enabling biothreats. Is it when an AI can independently solve a complex bioweapons design problem that no human without specialized training could? There is a need for clear frameworks to ensure AI developers are testing for biosecurity risks based on clear criteria before deployment.

Regulatory and Governance Gaps

The convergence of AI and biotechnology falls between regulatory silos. Existing biosecurity laws focus on dangerous pathogens and physical materials, while AI governance is still nascent. Overlapping jurisdictions (DoD, HHS, DHS, etc.) complicate a unified response. For example, who should regulate an AI that can design DNA? Should it be treated as a "munition" under export control, or as a software product, or as a public health issue? The pace of innovation in AI and biotech further strains regulatory agility. Internationally, there is an uneven patchwork of rules – what one country bans, another may allow in the name of research freedom, and AI models or bioexperiments can migrate accordingly. One proposal from experts is a future licensing regime for particularly high-risk AI-driven bio tools – akin to how nuclear technology is controlled – but only if and when such AI capabilities clearly emerge.

Escalation and Strategic Stability

On a geopolitical level, the AI-bio convergence could introduce new risks in deterrence and arms control. If states perceive that others are developing AI-optimized bioweapons or defenses, it could spur an arms race. Conversely, misinformation or AI-generated false alarms about biothreats could create confusion in crises. Maintaining transparency and communication internationally will be important to avoid worst-case assumptions.

Conclusion

The intersection of AI and biotechnology poses complex policy challenges. The risks are real but nuanced: we must neither be complacent nor hyperbolic. U.S. policymakers should aim to stay ahead of the curve by crafting flexible, forward-looking strategies that harness the benefits of bio and AI while minimizing misuse. This would include work to modernize biosecurity concepts, institute prevention safeguards that protect innovation, mandate and evolve the research on red-teaming AI systems for biosecurity risks, and improve early warning systems for monitoring and detection. Finally, the U.S. must work towards establishing a regime of oversight and regulation over high-risk AI systems that protects innovation and public safety, and continue to lead the world on international norms and standards on AI and biosecurity.

RISK MITIGATION IN BIO, AI, AND AT THE BIO × AI CONVERGENCE

DR. OLIVER STEPHENSON, FAS

As artificial intelligence capabilities expand into biological research and development, the intersection of these technologies creates emerging security challenges that are attracting increasing policy attention. Current mitigation tools—ranging from export controls to voluntary safety standards on AI models—offer partial protection but have gaps in coverage and enforcement. However, these protection methods are being developed in an environment where there remains debate about the exact nature of the risks and appropriate risk thresholds.

This brief examines why risk mitigation at the AIxBio nexus has become more of a pressing concern. It then discusses existing risk mitigation mechanisms, outlines current gaps, and provides questions to frame the conversation.

Why Mitigation Matters Now

The emphasis on addressing AI-biology convergence risks stems from two accelerating trends that are altering the threat landscape. First, biological capabilities are becoming increasingly accessible: DNA synthesis costs have dropped from around $10 per base pair in 2000 to less than $0.10 per base pair today.[8] While these declining costs have largely plateaued, some experts conjecture that this trend, in conjunction with a growing number of individuals with advanced synthetic biology skills, means that biological capabilities previously limited to nation-states or well-funded organizations are now within reach of smaller actors.

Simultaneously, AI systems are achieving unprecedented capabilities across a range of tasks, including biological domains. Foundation models can now predict many protein structures with near-experimental accuracy, design novel biological sequences, and propose synthesis pathways for complex molecules—tasks that previously required years of specialized training.[9]

While AI has many exciting applications in biology, researchers have raised concerns that this convergence could pose novel threats, and there have been increasing calls for governance and safeguards.[10] Studies by RAND and OpenAI published at the beginning of 2024 found at most a small uplift in biological risks posed by AI models,[11] and a study published in December 2024 stated that "existing studies around AI-related biorisk are nascent, often speculative in nature, or limited in terms of their methodological maturity and transparency."[12] However, in May 2025, leading AI company Anthropic activated additional safety precautions for their latest AI model, Claude Opus 4, due to its "continued improvements in CBRN-related knowledge and capabilities".[13]

[8] Carlson, Rob. How Fast Is The Energy Transition Going? That Depends On Where We Are Headed. (Synthesis.cc, 2026). https://www.synthesis.cc/synthesis

[9] Jumper, John, et al. Highly Accurate Protein Structure Prediction with AlphaFold. (Nature Publishing, 2021). https://www.nature.com/articles/s41586-021-03819-2

[10] Wang, Mengdi, et al. A Call for Built-in Biosecurity Safeguards for Generative AI Tools. (Nature Biotechnology, 2025). https://www.nature.com/articles/s41587-025-02650-8; Bloomfield, Doni, et al. AI and Biosecurity: The Need for Governance. (Science, 2024). https://www.science.org/doi/10.1126/science.adq1977

[11] Patwardhan, Tejal, et al. Building an early warning system for LLM-aided biological threat creation. (OpenAI, 2024). https://openai.com/index/building-an-early-warning-system-for-llm-aided-biological-threat-creation/; Mouton, Christopher, et al. The Operational Risks of AI in Large-Scale Biological Attacks. (RAND, 2024). https://www.rand.org/pubs/research_reports/RRA2977-2.html

[12] Peppin, Adrian, et al. The Reality of AI and Biorisk. (ArXiV, 2025). https://arxiv.org/abs/2412.01946

[13] Activating AI Safety Level 3 Protections. (Anthropic, 2025). https://www.anthropic.com/news/activating-asl3-protections

Current Tools in the AI and Bio Risk Mitigation Toolbox

Risk mitigation at the AI and Bio intersection currently relies on a range of tools spanning the technical, regulatory, institutional, and cultural. While AI and Bio mitigation tools have historically evolved along different tracks, these tools are converging along with the underlying technology.

Export controls and supply chain restrictions represent the most established mitigation approach, built on decades of experience controlling dual-use technologies. The Commerce Department's Entity List[14] now includes several biotechnology companies and research institutions, while the Bureau of Industry and Security has expanded dual-use export controls to include advanced biological equipment and materials.[15] These controls create bottlenecks in the global supply chain for sensitive biotechnology, forcing potential bad actors to either seek alternative sources or develop capabilities domestically—both costly and time-consuming approaches.

Screening and detection systems form a second major category of mitigation tools, operating primarily at the point where biological materials are synthesized or procured. The Department of Health and Human Services maintains the Federal Select Agent Program, which requires registration and oversight for research involving the most dangerous pathogens.[16] Meanwhile, industry players have developed voluntary screening protocols, for example, screening gene synthesis orders against databases of known pathogenic sequences,[17] with some additional guidelines and restrictions imposed by the federal government.[18]

Institutional and research governance—mechanisms like Institutional Review Boards (IRBs), Dual-Use Research of Concern (DURC) policies, and the P3CO (Potential Pandemic Pathogen Care and Oversight) framework[19] are designed to ensure that life sciences research—particularly with dual-use or high-risk potential—is subject to ethical and safety review. These institutional safeguards apply upstream, influencing what research is conducted and how. In AI contexts, similar oversight mechanisms are emerging in research labs and federal agencies, where model deployment or publication may trigger internal or external reviews (e.g., Anthropic's AI Safety Levels[20]).

AI safeguards represent the newest category of mitigation tools, designed specifically to address AI systems used in biological applications. These safeguards can be integrated at the AI model level, for example, removing potentially dangerous biological information from the training data, designing the model to refuse certain requests, watermarking model outputs, and unlearning harmful information.[21] Safeguards can also be at the system level, for example, flagging potentially harmful prompts for human review, and human red-teaming pre-deployment.

Information and publication controls aim to prevent the dissemination of sensitive research findings, instructions, or AI-generated outputs that could lower the barrier to misuse. Pre-publication review systems, controlled access to datasets, classification guidelines, and responsible disclosure practices are designed to mitigate risks. In AI, there has been increasing debate about the availability of open-weight models, especially in the life sciences, as once released, open-weight models cannot be withdrawn.[22]

[14] Entity List FAQs. (Bureau of Industry and Security, 2025). https://www.bis.gov/media/documents/entity-list-faqs.pdf

[15] Shepardson, David. US imposing new export controls on biotech equipment over China concerns. (Reuters, 2025). https://www.reuters.com/technology/us-imposing-new-export-controls-biotechnology-equipment-2025-01-15/

[16] 2024 Annual Report of the Federal Select Agent Program. (Centers for Disease Control, 2024). https://www.selectagents.gov/index.htm

[17] International Gene Synthesis Consortium. https://genesynthesisconsortium.org/

[18] 2024 OSTP Framework for Nucleic Acid Synthesis Screening. (Administration for Strategic Preparedness and Response, 2024). https://aspr.hhs.gov/S3/Pages/OSTP-Framework-for-Nucleic-Acid-Synthesis-Screening.aspx

[19] Enhanced Potential Pandemic Pathogen Oversight Framework. (Administration for Strategic Preparedness and Response, 2024). https://aspr.hhs.gov/S3/Pages/Enhanced-Potential-Pandemic-Pathogen-Oversight-Framework.aspx

[20] Activating AI Safety Level 3 Protections (q.v., 13).

[21] U.S. AI Safety Institute. Managing Misuse Risk for Dual-Use Foundation Models. (National Institute for Standards and Technology). https://doi.org/10.6028/NIST.AI.800-1.ipd; A Call for Built-in Biosecurity Safeguards for Generative AI Tools. (q.v., 10); Li, Nathaniel, et al. The WMDP Benchmark: Measuring and Reducing Malicious Use With Unlearning (ArXiv, 2024). https://arxiv.org/abs/2403.03218

[22] Kapoor, Sayash, et al. On the Societal Impact of Open Foundation Models. (ArXiv, 2024). https://arxiv.org/abs/2403.07918

International coordination and norms. Because advances in both biotechnology and AI diffuse rapidly across borders, durable risk mitigation depends on shared rules, transparency measures, and joint technical standards. Multilateral export-control regimes such as the Australia Group[23] already harmonize control lists for pathogens and dual-use equipment, while the Biological Weapons Convention (BWC)[24] provides a legal anchor for combating emerging biothreats. On the AI side, for a like the OECD AI Principles, the G7 "Hiroshima AI Process," and the EU-U.S. Trade and Technology Council are working on baseline expectations for model transparency, evaluation, and incident reporting.

Challenges and Gaps

The convergence of AI and biology is highlighting numerous gaps in the mitigation toolbox.

Export controls face inherent limitations in the AI-biology convergence space. Unlike traditional dual-use technologies that require physical components, many AI-enabled biological capabilities can be transmitted as pure information. A sophisticated AI model capable of designing novel pathogens can be shared instantly across borders through digital channels, making traditional export control mechanisms less effective.

Screening approaches are facing challenges as AI capabilities advance. Traditional sequence-based screening relies on comparing synthesis requests to databases of known harmful genetic material. However, AI systems may increasingly be able to design novel biological threats that share no sequence similarity with known pathogens while maintaining dangerous functionality—a capability that presents issues for existing detection frameworks.

AI safeguards are still nascent, with many of the fundamental techniques still being developed. AI developers often face challenging tradeoffs with dual-use capabilities, where restricting harmful model outputs can also cause harmless requests to be refused, limiting an AI model's utility. Additionally, while research has improved model security,[25] many model safeguards can be easily overcome through jailbreaking.

Information and publication controls relating to AI models with potentially dual-use biological capabilities are difficult to define. AI model capabilities are often not well understood at the point of release to the public, and it is still not clear how to translate outputs from AI benchmarking and red-teaming exercises into actual measures of risk. These concerns are heightened by the growing integration of large language models, biological design tools, and cloud labs in the biosciences, which are particularly challenging to evaluate for risk. These issues present particular challenges for the release of open-weight models, which can be enormously beneficial but cannot be restricted if unacceptably dangerous capabilities are discovered after release.[26]

International coordination and norms remain nascent with respect to potential AIxBio threats, with different norms around model evaluation, controls, and release between countries. Clashing approaches to regulation often make finding consensus difficult. Starkly different international regimes risk undermining controls imposed in any one country.

Beyond the specific gaps listed above, there remain three more fundamental challenges.

1. **A lack of widespread agreement** on the nature of the risks and appropriate mitigations. For example, researchers are still debating how the value of different forms of information provided by a language model could impact biological risks.

[23] The Australia Group: Fighting the spread of chemical and biological weapons. (Australian Department of Foreign Affairs and Trade, 2025). https://www.dfat.gov.au/publications/minisite/theaustraliagroupnet/site/en/index.html

[24] Biological Weapons. (United Nations Office for Disarmament Affairs). https://disarmament.unoda.org/biological-weapons/

[25] Sharma, Mrinank, et al. Constitutional Classifiers: Defending against Universal Jailbreaks across Thousands of Hours of Red Teaming. (ArXiv, 2025). https://arxiv.org/abs/2501.18837

[26] On the Societal Impact of Open Foundation Models. (q.v., 22).

2. **The problem of predicting future capabilities and risks** in the face of rapid technological change in both AI and biology. This speed means accurate predictions about future AI capabilities and risks are challenging, and regulations are hard to craft in a future-proof manner.
3. **A small talent pool at the AIxBio intersection.** Given the rapid progress in both fields, it is difficult to keep at the forefront of AI and Bio simultaneously, meaning the pool of people able to understand the risks, then develop and implement mitigations, is limited. .

LOOKING TO THE FUTURE AT THE BIO X AI RISK NEXUS

DR. YONG-BEE LIM, FAS

Bio x AI risks do not exist in a vacuum. The pace of technological evolution and maturation drives these risks, but key factors also include expert discourse and the way governments and societies frame the nexus of these issues. In this short paper, we present how these and other factors may influence risk at the intersection of biology and artificial intelligence, and identify important trends to consider in the future.

Ongoing Technology Evolution

Even three years ago,[27] few policymakers or the public understood the progress being made on artificial intelligence. Today, AI models are making rapid advancements in key areas such as reasoning and problem-solving, knowledge representation, planning and decision-making, and learning. Two key examples include:

- **AlphaFold**, an AI program developed by DeepMind, has consistently shown high levels of accuracy in predicting protein structures since it was first officially released in 2018 as AlphaFold 1.[28]
- **Large Language Models (LLMs)** are demonstrating aspects of expertise that match, or even exceed, those of PhD-level scientists through benchmarks such as the Virology Capabilities Test (VCT).[29]

Biology is also in the middle of a set of rapid advancements, particularly as it has interacted with other capabilities and technologies. These interactions are enabling scientists to generate new ways to work and process knowledge, including:

- **Cloud laboratories**, which are largely automated, and centralized laboratory facilities. Scientists can, to a certain extent, run entire experiments at a remote location with just their computer and a connection to the internet.[30]
- **Biofoundries**, which allow scientists to test large-scale designs and deploy microbes for everything from experiments and research to actual products that underpin the emerging bioeconomy.[31]

It is an axiom that these two accelerating technologies are examples of dual-use tools (that things may be used for both beneficial and harmful ends). Thus, the confluence of these two tools have the vast potential to exacerbate current biological weapons risks. The very tools and knowledge that inform drug discovery, delivery, and other mechanisms can be misappropriated to do great individual, community, and, in certain cases, large-scale harm.

Expert Discourse

There is no agreement among disparate communities as to the most likely risks of this nexus, as well as what risks should be prioritized. For example, some see the emergence of artificial general intelligence (AGI) as the ultimate end of humanity, while others welcome it with open arms. Some experts have divided AI risks into four categories (malicious use, AI race, organizational, and rogue AIs), but how likely and how these will manifest is

[27] Kalil, Tom. 2022 Bioautomation Challenge: Investing In Automating Protein Engineering. (Federation of American Scientists, 2022). https://fas.org/publication/2022-bioautomation-challenge-investing-in-automating-protein-engineering/

[28] Berke, Allison. Can't quite develop that dangerous pathogen? AI may soon be able to help. (Bulletin of the Atomic Scientists, 2023). https://thebulletin.org/2023/11/cant-quite-develop-that-dangerous-pathogen-ai-may-soon-be-able-to-help/

[29] Hendrycks, Dan & Hiscott, Laura. AIs Are Disseminating Expert-Level Virology Skills. (AI Frontiers, 2025). https://www.ai-frontiers.org/articles/ais-are-disseminating-expert-level-virology-skills

[30] Ireland, Tom. What are Cloud Labs? (Royal Society of Biology, 2022). https://thebiologist.rsb.org.uk/biologist-features/the-biologist-s-guide-to-cloud-labs

[31] Hillson, Nathan, et al. Building a Global Alliance of Biofoundries. (Nature Communications, 2019). https://www.nature.com/articles/s41467-019-10079-2

still an unanswered question as we look to the future.[32] Each may have specific implications for different aspects of biology.

A similar story exists in the biological security (or biosecurity) space. Some experts believe biological weapons are an inevitability, others believe that there are significant barriers to biological weapons research and development, and yet others see both sides and build hybrid perspectives.[33] This, amongst other challenges towards consensus, has led to differing views on the value of medical countermeasures, the impact of artificial intelligence, and how to address the dual-use nature of knowledge.

Further, none of these communities are static—rather, they continue to evolve, leading to changes in how we define, frame, and conceptualize biological risks. For example, the definition of biosecurity has evolved. The historic goal of biosecurity has been to ensure that only authorized individuals have access to the biology-related materials, equipment, knowledge, and facilities. These definitions have slowly been expanding, particularly as communities and countries merge concepts of biosecurity with those in adjacent areas such as health security and public health.[34] Some consider this a welcome change to the siloed status quo, while others highlight the challenges such an expansion raises.

Narrative Framing of Bio and AI

Various societies have touchpoints which shape the narrative of artificial intelligence - one cannot go into an AI meeting without at least one mention of SkyNet, for example.[35] A literature analysis of artificial intelligence highlighted four prominent AI narratives that are prominent in governmental and public discourse:[36]

- **The Competition Frame**: The world is engaged in an international AI "arms race", where two or more actors seek some form of superiority or strategic advantage, with AI poised as the game-changer.
- **The "Killer Robots" Frame**: Countries use autonomous systems to target and engage in combat without human supervision.
- **The "World Without Work" Frame**: AI replaces, rather than augments, human labor. There are many speculations about what may emerge in this world, with key concerns being reconfiguring our systems as needed to support such a dynamic, as well as potentially the loss of knowledge and skills over time.
- **The "Economic Gold Rush" Frame**: AI unleashes productivity and generates massive wealth for the global economy. This frame is often thought of as a paradigm shift from existing global economic structures to novel and disruptive ones.

The narrative around biology is a challenging one to navigate. Life sciences research, in particular, has been experiencing a large amount of scrutiny due to narratives emerging from the COVID-19 pandemic. This has led to what feels like a polarized world in biology, which creates alignment challenges as our global community faces a potential world of increasing biological risks. Key narratives at this time include:

- **Biology as compounding existential risks**: From narratives of risky experimentation to the impacts of COVID-19, the decisions and practices surrounding life sciences research have been increasingly probed.

[32] Hendrycks, Dan, et al. An Overview of Catastrophic AI Risks. (ArXiV, 2023). https://arxiv.org/abs/2306.12001

[33] Koblentz, Gregory. Predicting Peril or the Peril of Prediction? Assessing the Risk of CBRN Terrorism. (Terrorism and Political Violence, 2011). https://www.tandfonline.com/doi/abs/10.1080/09546553.2011.575487

[34] DiEulis, Diane. Bolstering Biosecurity Amid the Biotechnology Revolution. (Orbis, 2024). https://www.sciencedirect.com/science/article/abs/pii/S0030438724000255

[35] SkyNet (Terminator). https://en.wikipedia.org/wiki/Skynet_(Terminator)

[36] Imbrie, Andrew, et al. Contending Frames: Evaluating Rhetorical Dynamics in AI. (Center for Security and Emerging Technology, 2021). https://cset.georgetown.edu/publication/contending-frames/

- **Biology as the disruptive green technology of the future**: Some see biology as a greener alternative to current manufacturing practices, such as those that largely use chemical-based processes.
- **Biology as a major economic engine**: As the recent report from the National Security Commission on Emerging Biotechnology (NSCEB) notes,[37] using biology as the basis for innovation can provide enormous benefits, including its applications in defense, healthcare, agriculture, energy, and manufacturing.

Considerations for the Future

How we frame and mitigate risks from biology, AI, and their convergence will be a constant evolution—one that will be hard to predict, and will likely generate similar narratives that echo from the past. Three key considerations are provided below:

Biology, AI, and their convergence are all value-neutral tools. The outcomes from these technologies are determined by intent: how people intend to use these products for beneficial, nefarious, or other purposes in between the extreme ends of this spectrum. The complexity, particularly when it comes to AI, is that human intention is not the only determining factor in this construct. Rather, loss of control over AI systems implies these tools themselves may operate with intent - something that is of concern that should be explored in the future.

As we consider the risks, we must understand there is a significant relationship between technological hype and feast-and-famine cycles. As new technologies emerge, many experts attempt to predict the specific offense/defense implications of said technologies. Given the speculative nature of these assessments, we need to balance the emergence of probable risks and their scale of impact with ensuring sustained efforts in areas such as biosurveillance and AI safety assessments.

Intent is as key as capability in future discussions. Technological barriers to the misuse of biology and AI are shifting based on where people live and what capabilities they have access to. However, the intent to use technology for nefarious purposes is the trigger for misappropriating capabilities. Therefore, it is important to consider not just how capabilities are advancing, but how military doctrine, strategy, norms, and societies are evolving, which may make such nefarious uses more or less acceptable.

Conclusion

The risks from biology, artificial intelligence, and their convergence raise significant concerns in a geopolitically heated and complex world. As these technologies advance and experts grapple with the best path forward, cooperation is the key to unlocking a safer, more secure world from global catastrophic risks: cooperation that we are looking to build and expand on through our work at the Federation of American Scientists in partnership with the Future of Life Institute.

[37] Federation of American Scientists (FAS). Position On National Security Commission On Emerging Biotechnology Final Report: Charting The Future Of Biotechnology. (FAS, 2025). https://fas.org/publication/charting-the-future-of-biotechnology/

Interested to learn more about our AIxGlobal Risk Nexus Series?

Visit FAS.org to learn more about our upcoming events, publications and Global Summit 2026.

FAS FEDERATION OF AMERICAN SCIENTISTS

ABOUT THE FEDERATION OF AMERICAN SCIENTISTS

The Federation of American Scientists is dedicated to democratizing the policymaking process by working with new and expert voices across the science and technology community, helping to develop actionable policies that can improve the lives of all Americans. For more about the Federation of American Scientists, visit **FAS.org.**