Research Report

ANTON SHENK, MATT CHESSEN, BARBARA DEL CASTELLO, GREGORY SMITH,
RICHARD S. GIRVEN

Infinite Potential—Insights from the Viral Uplift Scenario

After-Action Report from a Sequence of Day After Artificial General Intelligence Exercises

For more information on this publication, visit **www.rand.org/t/RRA4727-1**.

Published by the RAND Corporation, Santa Monica, Calif.

RR-A4727-1

About This Report

To understand how the United States should prepare for and respond to potential artificial intelligence (AI) and artificial general intelligence (AGI) developments, the RAND Geopolitics of AGI Initiative conducts Day After AGI exercises using the RAND Infinite Potential platform. Each exercise simulates a National Security Council Principals Committee convening to recommend a U.S. government response to frontier AI developments. Participants confront a scenario that poses both an acute crisis for U.S. national or economic security and a signpost on the path to a transformative AI future. Guided by a simulated National Security Advisor, they role-play as Principals Committee members—diagnosing the scenario's implications for U.S. national security, deliberating over courses of action prepared by their staffs, and recommending a path forward to the President of the United States.

This report draws on ten runs of the Viral Uplift scenario, which unfolded across two turns. In the first, participants confronted a global pandemic triggered by students' accidental release of an AI-engineered virus. In the second, participants faced intelligence revealing that multiple state and nonstate actors were racing to acquire similar AI-enabled biological weapon capabilities. The scenario was designed to surface policy options for reducing both the risk and impact of AI-model misuse in generating novel biological threats. Across all ten runs, the exercises illuminated critical decision points, capability gaps, and institutional challenges that may emerge as AI reshapes the strategic landscape—enabling policymakers to better prepare for an uncertain but potentially transformative technological future.

This report is intended for policymakers, AI policy researchers, and members of the broader public seeking to understand potential AI crises and the analytic capabilities that may be necessary to respond to them.

Center for the Geopolitics of Artificial General Intelligence

RAND Global and Emerging Risks is a division of RAND that delivers rigorous and objective public policy research on the most consequential challenges to civilization and global security. This work was undertaken by the division's Center for the Geopolitics of Artificial General Intelligence (AGI), which is committed to helping decisionmakers understand, anticipate, and prepare to navigate the national security and geopolitical implications of AGI. The center convenes leading technologists, strategists, economists, political scientists, and outside experts to consider the feasibility and effectiveness of prospective AGI-enabled capabilities; the domestic and international implications of their use; and the strategies and policies that governments, businesses, and civil society could adopt to respond to new realities. For more information, visit http://www.rand.org/geopolitics-of-agi.

Funding

This effort was independently initiated and conducted within the Center for the Geopolitics of Artificial General Intelligence using income from operations and gifts from RAND supporters, including philanthropic gifts made or recommended by DALHAP Investments Ltd., Ergo Impact, Founders Pledge, Charlottes och Fredriks Stiftelse, Good Ventures, Longview, and Coefficient Giving. RAND donors and grantors have no influence over research findings or recommendations.

Acknowledgments

We would like to thank the Geopolitics of AGI Center leadership and, in particular, Joel Predd and Emma Borden for their invaluable assistance in executing the exercises summarized in this document. We would also like to thank the exercise participants, without whom this report would not be possible.

Contents

About This Report iii
Figure and Tables vi

Infinite Potential—Insights from the Viral Uplift Scenario 1
Applying Gaming to the Problem of Artificial General Intelligence 2
Exercise Structure 4
Two-Turn Structure 4
Setting and Data Collection 5
Scenario Summary and Research Objectives 5
Exercises and Participants 7
Key Issues 8
Capabilities and Playbooks to Enhance U.S. National Preparedness 10
Next Steps 14

Appendix 15
Capabilities and Response Playbooks 15

References 18
About the Authors 19

Figure and Tables

Figures

Figure 1. Game Play and Analysis Process 5

Tables

Table 1. Scenario Summary 6
Table 2. Participant Summary 7
Table A.1. Summary of Proposed Capabilities 15
Table A.2. Summary of Proposed Response Playbooks 16

Infinite Potential—Insights from the Viral Uplift Scenario

In this report, we present an analysis of ten iterations of the Viral Uplift scenario, in which participants role-playing National Security Council Principals Committee members responded to artificial intelligence (AI)–enabled biological (AI-bio) crises. The scenario unfolded across two escalating phases: Participants first confronted an accidental pandemic caused by biological weapons that nonstate actors created using advanced AI tools, then faced intelligence revealing that multiple governments were racing to acquire similar capabilities.

Across ten exercises, 119 participants—drawn from senior government officials, RAND analysts, and subject-matter experts—surfaced five critical patterns concerning both the challenges posed by AI-enabled biological threats and the policy options for countering them:

1. **Governance strategies divided sharply between restricting AI capabilities and targeting dangerous actors.** Participants consistently debated whether to focus on controlling access to advanced AI systems or on disrupting the actors who might misuse them. All groups acknowledged the need for both approaches, but fundamental disagreements persisted over where to place emphasis. Some argued that restricting distribution of biologically capable AI models was essential to reducing risk; others countered that such restrictions were technically infeasible, given the proliferation of open-source models.
2. **Resilience, not containment, must anchor U.S. preparedness for and response to AI-enabled biological threats.** Across all ten exercises, no group identified a credible path to put the genie back in the bottle for dual-use AI capabilities. Participants instead pivoted toward building defensive capacity—expanding medical countermeasure (MCM) production, forging public-private partnerships for rapid threat response, and developing technical mechanisms to detect and counter AI-generated pathogens.
3. **China policy exposed deep tensions between cooperation on shared biological threats and competition over AI leadership.** Every exercise surfaced debate over engaging Beijing. Participants who assessed that China would prioritize the shared global risk of AI-enabled pandemics advocated for scientific cooperation and joint governance frameworks. Those who argued that China would exploit the crisis to advance its own AI capabilities favored unilateral action and allied coordination. No group reconciled these competing imperatives.
4. **Existing legal authorities and regulatory frameworks are inadequate for AI-enabled biological crises.** In all ten exercises, participants identified critical gaps in existing authorities for managing AI-enabled threats. Groups consistently concluded that current dual-use research oversight mechanisms—designed for traditional biotechnology—cannot keep pace with the speed and scale of AI-enabled biological capabilities. The groups called for new

emergency powers, reformed international frameworks, and enhanced authorities to compel private-sector cooperation.

5. **Enhanced biological surveillance and early warning systems emerged as the highest-priority capability needs.** Every exercise highlighted the critical importance of detecting AI-engineered pathogens before they spread widely. Participants consistently prioritized investments in next-generation sequencing, AI-powered anomaly detection, and real-time data sharing—emphasizing that traditional disease surveillance lacks the speed to counter deliberately engineered threats.

Applying Gaming to the Problem of Artificial General Intelligence

AI has become a topic of interest and concern within the U.S. policymaking community. There has been increasing interest in artificial general intelligence (AGI) and its potential to affect areas critical to national security, from economic growth to military power.[1] However, while AGI's potential impact is increasingly recognized, this technology remains hypothetical, and exactly how to define this technology is contested. Some define AGI as an AI that could teach technology that could perform every intellectual task a human could perform, while others focus on the ability for such an AI to rapidly reach superhuman levels of capability. The world has not yet seen an "AGI" and does not fully agree on what AGI even is.[2] This uncertainty, coupled with the potential size of the impacts that AGI might have, means that much of the response to AGI may occur only once the technology, or technologies like it, has been developed and deployed.

But that does not mean that analysis of AGI's national security implications is not important. Instead, analysis of AGI's impacts must occur under conditions of uncertainty. One way to clarify and explore this uncertainty is through gaming, which can be a useful tool for exploring the relationships within a policy problem when those relationships are unknown and/or when the factors important to that problem are not yet fully understood.[3] Games and exercises allow researchers to explore these uncertain problems in collaboration with their participants, drawing on their decisions and discussions to identify how a problem is structured, what factors will be particularly important for resolving or managing it, and the strength of the relationships.

The Day After AGI exercises within the Infinite Potential platform were launched to analyze an element of the policy problem AGI presents; the potential impacts of AGI are so large, yet so uncertain in how specifically they will manifest, that it is unlikely that the U.S. government will be able to address the challenges of AGI primarily through preemptive policymaking. Rather, it is more likely that U.S. policy with regard to AGI will evolve through crises and specific incidents that demand a response. However, it is still possible to prepare for crises through contingency planning and

---

[1] Jim Mitre and Joel B. Predd, *Artificial General Intelligence's Five Hard National Security Problems*, RAND Corporation, PE-A3691-4, February 2025.

[2] For an exploration of different ways to define AI capabilities, see Meredith Ringel Morris, Jascha Sohl-Dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet, and Shane Legg, "Levels of AGI for Operationalizing Progress on the Path to AGI," *Proceedings of the 41st International Conference on Machine Learning*, PMLR 235, 2024.

[3] Edward Parson, "What Can You Learn from a Game?" in Ralph L. Keeney, Richard J. Zeckhauser, and James K. Sebenius, eds., *Wise Choices: Decisions, Games, and Negotiations*, Harvard Business School Press, 1996.

preparation so that if and when an AGI or merely an advanced AI crisis occurs, the United States has greater capability to respond and bring the crisis to a conclusion.

This series of games and exercises was designed to answer three key questions about how the United States could prepare for the uncertain but potentially significant impacts of AGI and build knowledge ahead of such crises:

1. What key issues were participants in these exercises trying to make sense of in the face of AI or AGI-related crises, and what information did they seek to help them make judgments about those issues?
2. What capabilities did participants wish they had or wish to develop in response to the crisis?
3. What playbooks did participants identify need to be written and validated to adequately respond to the crisis presented to them in the exercise?

Each of these questions implies certain actions that may be worth taking now, in advance of a similar AGI crisis, should one occur. The key issues participants identified, the judgments they sought to make, and the questions they asked imply areas of uncertainty that could be clarified to answer the questions in advance of any crisis. In turn, the lists of capabilities and playbooks that participants desired during the exercise imply areas for investment and development by the U.S. government or other actors, either now or in the future, when such capabilities became more necessary.

To answer these questions, we designed a series of scenario-based exercises to confront participants with different potential AGI-related crises. Each scenario presented participants with different AI capabilities and their impacts on the world, ranging across multiple levels of AGI-related progress and capability, depending on the problem being presented to the participants.[4] By working through these crises in an exercise, participants were prompted to think about crisis response options that could explicitly or implicitly identify the issues, capabilities, and playbooks that could help resolve the crises if they ever occurred in the real world.

The exercise structure (described in more detail later) was designed to be easy to play; it would be run in two hours with no preparation required of participants so that the exercise could be run multiple times with diverse audiences to gather input on potential ways to respond to crises from a broader set of participants. Breadth of participants was considered particularly desirable because of the uncertainty surrounding AGI and the wide variety of implications it might have. Under such uncertainty, leveraging different experiences from different audiences was considered valuable for identifying the set of issues, capabilities, and playbooks that might address the scenario at hand.

Results of scenarios were compiled to identify common issues, playbooks, and capabilities that participants discussed during the runs. The results of these runs, captured by dedicated notetakers, were then compiled into after-action reports, such as this document, that discuss participant discussion, actions, and behavior across multiple exercises. This after-action report summarizes the results of ten runs of the Viral Uplift scenario.

[4] In the future, additional analysis will identify trends across a sufficiently large number of runs of an individual scenario to permit judgments about the best policies for the U.S. government and other actors to pursue. Analysis may also be written identifying trends across different scenarios.

Exercise Structure

The research team designed scenario-based exercises to confront participants with plausible AI/AGI-related crises spanning multiple capability levels and domains of impact. These scenarios aim to stimulate crisis-response thinking that would reveal, either explicitly or implicitly, the issues, capabilities, and playbooks required for effective real-world response.

Infinite Potential exercises are two-hour tabletop exercises patterned after the Day After format, with ten to 16 participants led by a facilitator, whose job is to elicit responses and discussion from the group. The exercise structure prioritizes accessibility and breadth—each session runs two hours with no advance preparation, enabling multiple iterations with diverse participant groups. This design philosophy reflects AGI's broad uncertainty: Given the technology's unclear implications across domains, incorporating varied professional experiences and perspectives enhances the likelihood of identifying comprehensive requirements for crisis response.

Participants were selected based on their experience, institutional knowledge, and expertise. In the ten runs summarized here, participants took on the roles of members of the National Security Council Principals Committee, with the facilitator playing the role of the National Security Advisor convening the meeting. Participants were asked to engage in light role-play of their assigned roles, representing the equities of their assigned positions but not being bound to them or the model of any former or current occupant of that position. The participants were asked to diagnose the implications of the scenario for U.S. national security, discuss courses of action presented by their staffs, and recommend a path ahead for the President.

However, the participants were not asked to assume that any particular U.S. administration would be in office at the time of the exercise scenario. Instead, they were asked to consider the scenario from the perspective of an administration confronting this issue, drawing on their historical knowledge of U.S. national security.

Two-Turn Structure

Each exercise unfolds over two turns that escalate the crisis:

- **Turn 1** begins with participants receiving intelligence and open-source information about an emerging AGI-related crisis. A subject-matter expert briefs the team on the unfolding situation, after which participants can ask clarifying questions about the events. The facilitator then guides the group through structured discussions of policy objectives, available response options, and the potential ramifications of different courses of action.
- **Turn 2** progresses the scenario—typically an escalation that worsens the crisis, although the situation sometimes develops along an unexpected trajectory. Following another expert briefing on these accelerated developments, participants again explore policy issues and response options with their enhanced understanding of the crisis dynamics. The turn

concludes with a retrospective evaluation where participants assess their initial responses and strategies.[5]

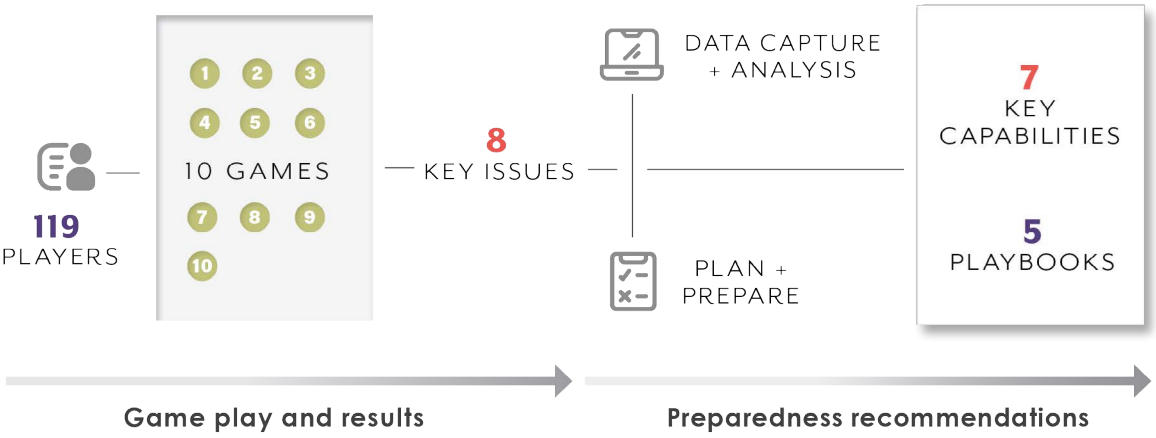
Setting and Data Collection

To ground the crisis in realistic near-term conditions, gameplay is set in a hypothetical future of 12 to 18 months from the current date. This time frame was selected because AI's rapid pace of improvement motivates exploring what a potential AI crisis could look like in the short term. Participants are instructed to assume the technological and geopolitical landscape will not have changed significantly, with the realities of the current year persisting in the exercise.

During each exercise, notetakers captured participant discussions—which are then reviewed and compiled to identify recurring issues, capabilities, and playbooks across groups. This report analyzes results from ten iterations of the Viral Uplift scenario, examining how participants approached AI-enabled biosecurity crises and identifying the measures they deemed most critical for effective response.

Figure 1 displays the process described for the Viral Uplift scenario, from the gathering of players and running the exercise through the identification of key issues, capabilities, and playbooks from the exercises.

Figure 1. Game Play and Analysis Process



Scenario Summary and Research Objectives

In each exercise using the Viral Uplift scenario, participants engaged in two turns (see Table 1). In Turn 1, students affiliated with a radical student group used a jailbroken open-source AI model alongside an AI biological tool to engineer a modified SARS-CoV-2 (COVID) variant. The AI enabled the group to overcome technical barriers in viral engineering by interfacing with the biological

[5] In some iterations, teams also engage in *backcasting* exercises to identify preventive measures or mitigation strategies that could have been implemented before the crisis began.

tool and providing step-by-step guidance. The resulting neo-COVID virus was accidentally released, causing a global pandemic with relatively low fatality rates but demonstrating the proliferation potential of AI-enabled biological weapon capabilities.

In Turn 2, intelligence revealed multiple state and nonstate actors attempting to acquire similar AI-enabled capabilities for developing biological weapons. These threats included a foreign government pursuing an offensive biological weapon program against another nation, a terrorist group seeking to engineer ricin-producing *Escherichia coli* bacteria while offering to sell the information to other terrorist organizations, and the discovery of a potentially dangerous new biological AI tool at a U.S. biotechnology company.

Table 1. Scenario Summary

Turn	Time	Scenario Synopsis	Decisions
Turn 1	Late 2025	Nonstate actors create AI-enabled biological weapons, resulting in the accidental release of an engineered virus that becomes a pandemic.	How should the U.S. government respond to proliferation of biological weapon capabilities through AI?
Turn 2	+3 months	Additional state and nonstate actors gain access to AI capable of supporting the creation of bioweapons. In addition, AI technology continues to advance, and more-capable AI that can provide greater support for bioweapon creation begins to proliferate.	How should the U.S. government respond to escalating AI-enabled biological threats from multiple actors?

Scenario Research Objectives

This scenario was developed to explore policy options for addressing a common threat vector in existing literature about advanced AI: the AI-enabled creation of dangerous pathogens by state and nonstate actors. This threat is regularly cited as a near-term risk in current policy discussions, although the specific threats such tools might pose remain largely undefined.[6] Given this uncertainty and the potential severity of biological crises, we developed this scenario to explore how the U.S. government might respond to AI-enabled biosecurity threats and to identify preparatory actions that could reduce the impacts of future crises involving these emerging capabilities.

The AI tools depicted in the scenario are capable of human-level performance across most elements of biological science, including the design and synthesis of novel pathogens. However, the scenario does not specify whether these tools constitute AGI in the broadest sense of that term; rather, they represent highly capable AI systems that have achieved transformative impact in the biological sciences. The scenario invites players to consider whether and how such capabilities might transfer to other domains, although it does not specify such cross-domain capabilities.

[6] John Halstead, "Managing Risks from AI-Enabled Biological Tools," *GovAI* blog, August 5, 2024.

Exercises and Participants

Ten exercises were conducted between April and December 2025. Three exercises contained a mix of current or former senior principals playing Principals Committee roles with which they had direct experience (see Table 2). Four exercises involved mostly or all RAND analysts and researchers, with some participants from other research organizations. Three exercises consisted of current U.S. government staff members who worked directly on biological research and biosecurity issues.

Table 2. Participant Summary

	Exercises									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Current or former U.S. government principals	0	1	12	1	0	0	0	0	0	0
Current or former U.S. government staff	2	0	0	3	0	0	10	15	15	9
Senior RAND or other analysts	7	11	0	12	1	5	6	0	0	0
Junior RAND or other analysts	8	0	0	0	8	7	0	0	0	0

The exercises incorporated three distinct participant categories to capture the diverse perspectives required for effective crisis response. Current government officials included both political appointees and civil servants holding midlevel and senior positions responsible for policy formulation and implementation across their agencies. The majority were career civil servants, supplemented by military officers from the Department of War and representatives of the Department of State, Department of the Treasury, Department of Commerce, and the intelligence community.[7] Former government officials brought seasoned expertise from across multiple administrations, including former deputy secretaries and senior political appointees with substantial decisionmaking authority. These participants typically possessed more than five years of experience in senior strategy, planning, and implementation roles for high-level national security and technology policy decisions. RAND and external analysts contributed specialized knowledge of AI, AI policy, and national security implications. External participants were drawn from leading academic institutions and think tanks recognized for expertise in emerging technology policy.

[7] The Department of War is designated the Department of Defense under Public Law 81-216, National Security Act Amendments of 1949.

Key Issues

Participants' assessments of threats to U.S. national security—and the policy responses they recommended—consistently hinged on how they judged a set of key issues. RAND analysts who observed the exercises and reviewed transcripts identified these issues as the most critical in affecting participants' judgments. Most typically, participants agreed on these issues' importance but disagreed, often sharply, on how to resolve them—reflective of genuine underlying uncertainties. Enabling better judgments on these issues implies different collection and analysis requirements for the U.S. government.

Key Issue 1. AI as Force Multiplier Versus Game Changer for Biological Threats

Participants in all exercises debated the degree to which AI genuinely enables the development of biological weapons, as opposed to simply lowering barriers for existing capabilities. Key uncertainties for participants included understanding how effectively general-purpose AI models could interface with specialized biological tools and how quickly such capabilities would advance, as well as exactly how much uplift a nonexpert would receive from working with such a model. In six of the ten exercises, participants were also uncertain about how rapidly AI models would overcome remaining bottlenecks for gene synthesis. For example, in one exercise, participants concluded that the rate of improvement demonstrated in the scenario suggested a very short timeline for continued improvement. No consensus emerged on the speed of continued model improvement, but participants agreed that resolving this question would be critical for determining how quickly preparedness measures and policy responses would need to be implemented once a risk materialized.

Key Issue 2. Adequacy of Attribution and Legal Frameworks

Participants in all exercises identified gaps in existing authorities (e.g., Biological Weapons Convention, domestic counterterrorism laws) for responding to AI-enabled biological threats. Questions emerged about when possession of dangerous AI models constitutes material support for terrorism under current laws and whether possession should count as support, given the scenario presented to players. Participants also questioned whether the U.S. government could attribute biological attacks to specific AI tools or developers and whether there were sufficient legal authorities to take down such dangerous models, should they be detected. In all exercises, participants concluded that current authorities were insufficient to resolve these uncertainties and would be difficult to exercise in a crisis.

Key Issue 3. Chase the Code Versus Chase the Actor

In every exercise, participants discussed how to balance between attempting to restrict AI models and focusing on disrupting specific threat actors. Participants agreed that control of open-source models is extremely difficult, leading many to favor actor-focused approaches or traditional biosecurity

measures designed to keep these actors from accessing biological materials and laboratory equipment. However, this raised questions about surveillance authorities, the need for international cooperation, and the sustainability of actor-focused approaches as capabilities proliferate.

Participants also suggested in two exercises that there might be value in controlling key virology-related datasets and the compute that would be required to run a model. While participants agreed that controlling such inputs to AI models would not fully resolve the issues that open-source models present, they stated that controls on data and chips could be useful as part of a multilayered approach to make it harder for threat actors to use AI for dangerous purposes.

Key Issue 4. Offense-Defense Balance in AI-Bio

In every exercise played, participants debated among themselves whether AI would provide offensive or defensive advantages in biosecurity. Key uncertainties included what role AI could play in improving biosurveillance and early warning systems and whether defensive AI applications require the same dangerous capabilities as offensive ones or could be developed independently.

Key Issue 5. Cooperate Versus Compete with Beijing

The global nature of AI development was a topic of discussion for participants in all exercises. Participants stated that effective preparedness and response would require international cooperation but questioned its feasibility, given the potential for conflicts between major nations. Participants were always divided on whether the United States should seek cooperation with China on managing risks from open-source models.

In two exercises, participants focused on how best to position the United States as a responsible steward of global health and safety. Players in these two exercises agreed that it would be difficult to comprehensively protect global health against an increase in biothreats without some degree of cooperation with China and a broader framework for engaging responsibly with the rest of the world but reached no consensus on how best to achieve either objective. Unresolved issues remained surrounding what sort of agreement could be acceptable to both the United States and Beijing and whether some form of global governance regime could be sufficiently agile to respond to the risks the scenario presented.

Key Issue 6. Escalation and Deterrence Thresholds

In eight exercises, participants debated response thresholds, particularly for state actors developing AI-enabled biological weapons. Participants in all exercises debated what level of threat would justify kinetic military responses and how the U.S. government could deter state and nonstate actors without provoking broader conflict.

Key Issue 7. Security Versus Scientific and Commercial Progress

In seven of the exercises, participants debated how to balance security against potential negative impacts on beneficial research, open-source model development, scientific collaboration, and the development of the U.S. biotechnology industry. Participants in these exercises were uncertain to what extent restrictions on AI-bio tools would impede legitimate research and commercial development of valuable treatments. In each of these exercises, participants sought to find ways to distinguish beneficial from harmful applications of dual-use technologies but found it difficult to draw a clear line. Participants ended up observing that the most useful biotechnology would also be the most dangerous.

Key Issue 8. Crisis Speed Versus Democratic Deliberation

Participants in five of the exercises voiced concern over the compressed timelines of AI-enabled threats and the deliberative processes essential to democratic governance. Participants said that normal U.S. government decisionmaking timelines might be inadequate for AI-related emergencies and were concerned that the U.S. bureaucracy might be too slow to respond to the threat presented in the scenario. Players in three exercises suggested that a U.S. government leader and lead federal agency would need to be empowered to respond to enhanced biological risks and take additional action to prevent them in the future but did not reach consensus on where within the U.S. bureaucracy such a leader should be positioned or what powers that leader would require.

Capabilities and Playbooks to Enhance U.S. National Preparedness

The key issues arising in the scenario suggest a set of capabilities and playbooks that could be developed to improve U.S. government preparedness for advanced AI. The lists in this section are not exhaustive of all capabilities that may be worthy of investment but were outcomes of these exercises.[8]

Capabilities

Capability 1: AI-Bio Risk Assessments and Technical Evaluation

In all exercises, participants raised the need for enhanced technical assessments of risk and stated that they were uncertain about the scale of the genuine enabling effects of AI on biological weapon development. All participants found the U.S. government's current inability to provide authoritative technical assessments to be problematic because of the national security nature of the threat. Also, in all exercises, participants advocated for dedicated technical teams capable of evaluating predeployment AI model biological capabilities, partnerships with leading AI and biotechnology companies for threat assessments, and standardized evaluation frameworks for AI-biological risk to fill this critical intelligence gap. Risk assessments, which seek understanding of exactly how AI-enabled biological

[8] The appendix summarizes the capabilities and playbooks that participants identified.

tools would be used in increasing biological risk, could then be developed by biosecurity experts. Technical evaluations would better inform the results of these assessments.

Capability 2: AI Model Proliferation Monitoring and Analysis

Participants across all exercises agreed that direct control of open-source AI models is extremely difficult but, nevertheless, advocated for enhanced monitoring capabilities to track dangerous model proliferation, should specific models be identified as high risk (either from incidents or from evaluations). Participants considered tracking the development of concerning AI-bio models to be a critical intelligence requirement and suggested it would allow either disrupting the distribution of problematic models or, at least, assessing their capabilities and use. Participants in all exercises also emphasized the need for capabilities to characterize dangerous models, monitoring systems for model downloads and use, and honeypots to identify malicious use. Participants in multiple exercises also noted that these monitoring and disruption capabilities would primarily be effective against models hosted on cloud infrastructure, noting that the proliferation of open-source models that can be downloaded and run locally poses a fundamental challenge to model-focused containment strategies.

Capability 3: Enhanced Biological Surveillance and Early Warning Systems

Participants in all exercises identified a need for dramatically improved biosurveillance, as AI-enabled threats could emerge and spread rapidly. Participants consistently emphasized that traditional approaches would be inadequate for AI-accelerated threats, which might be engineered to evade conventional detection. Participants raised several potential methods to enhance biosurveillance across exercises, including developing pathogen-agnostic wastewater surveillance systems, AI-enabled genomic sequencing and analysis, and integrated biological and cyber threat monitoring.

Capability 4: Rapid Response Biotechnology Development for Medical Countermeasures

Participants in all exercises stated that the potential for rapid, successive AI-enabled biological attacks meant that traditional MCM development would be inadequate to counter the biological threat presented in the scenario. Participants in all exercises stated that, if AI can accelerate development of biological weapons, defensive capabilities would need to achieve similar speed advantages. Participants suggested different capabilities under this umbrella in different exercises, with suggestions including AI-accelerated vaccine and therapeutic development platforms, flexible manufacturing capabilities for rapid scale-up, faster diagnostics, and other broad-spectrum MCMs for unknown biological threats that could be developed faster than traditional identification and response options would allow. Some participants acknowledged that broad-spectrum countermeasures for unknown biological threats may be very difficult to produce, but argued that the threat environment nonetheless demanded investment in this direction.

Participants in three exercises went deeper on this topic and focused on the need to better develop the domestic MCM supply chain. They pointed to shortages during the COVID-19 pandemic, when MCMs were often in short supply and could not be produced domestically, as demonstrating that

domestic capacity would need to be urgently developed for MCMs if AI's growing biological capability made pandemics more likely.

Capability 5: International Intelligence and Cooperation Framework

Participants in all exercises stated that effectively preparing for and responding to AI-bio threats requires unprecedented global cooperation, yet current mechanisms are inadequate for the speed and complexity of these threats. Across exercises, participants identified a need for intelligence-sharing mechanisms on AI-bio threats, diplomatic engagements on AI-bio governance, and enhanced monitoring of foreign AI-bio development programs and threat actors. Participants in two exercises went into more detail on this capability, identifying that all-source fusion of intelligence from across the intelligence community and the collection of new intelligence on potential biological threats from AI would be necessary to ensure that dangerous nonstate actors could be identified before they deployed a biological weapon.

Capability 6: Enhanced Counterintelligence for AI-Bio Convergence

Participants in seven exercises highlighted the vulnerability of U.S. AI and biotechnology companies to foreign intelligence collection and criminal acquisition, particularly given the dual-use nature of AI-bio technologies. Participants advocated for different solutions to this issue in different exercises, including enhanced monitoring of foreign intelligence activities targeting U.S. AI and biotechnology companies, improved vetting and monitoring of researchers with access to dual-use AI-bio tools, and capabilities to detect and disrupt foreign acquisition of sensitive AI-bio technologies.

Capability 7: Legal and Regulatory Preparedness and Response Framework

Participants in five exercises focused on the significant gaps in existing legal authorities for preparing for and responding to AI-enabled biological threats. These groups of participants stated that the compressed timeline of AI-enabled threats requires prepositioned legal authorities that can be activated rapidly during crises, particularly if the crisis involves both foreign and domestic threat actors. Participants noted that some proposed authorities, particularly those involving model takedown or information operations against domestic servers, would need to be carefully designed to be legally permissible under the Constitution. These participants did not believe the current frameworks could enable such a rapid response.

Playbooks

Playbook 1: AI-Bio Crisis Response Protocol

While participants did not definitively assess the technical feasibility of various AI-bio threats in the exercises run, they still developed responses for managing such crises. Participants who emphasized the urgency of AI-enabled threats favored more aggressive coordinated responses across government agencies, while those who questioned the technical assessments discussed more measured approaches focused on verification and deliberation. Across all exercises, participants agreed on the need for such procedures as rapid threat assessment and characterization with standardized AI model

evaluation protocols for biological risk; coordination between health, security, and technology agencies; communication protocols for public messaging and international engagement; and escalation criteria for different levels of response.

Playbook 2: Model Containment and Disruption Operations

With the challenge of controlling AI model proliferation being a concern throughout the exercise, participants in all exercises sought to develop approaches for managing dangerous models while acknowledging the limitations of direct control. In all exercises, there was a split in the preferred approach of participants. One group of participants called for aggressive takedown efforts in coordination with digital platforms, while others emphasized the intelligence value of monitoring model use. Participants agreed that some form of model-focused response capability was necessary, even if some stated they did not believe efforts targeting models would be sufficient alone.

The specific components of this playbook differed from exercise to exercise and included rapid identification and analysis of dangerous AI models, poisoning of certain dangerous datasets, coordination with platforms and cloud providers for model removal, information operations to disrupt threat actor access to dangerous models, and legal processes for compelling model removal or access restriction. Participants also noted that new authorities would likely be required, especially for takedown or information operation efforts. Players in about half of the exercises raised the concern that model containment and takedown measures apply primarily to cloud-hosted models and are of limited effectiveness against open-source models that have been downloaded and are run locally, a limitation that participants described as a fundamental constraint on model-focused approaches.

Playbook 3: International Cooperation and Diplomatic Engagement

Participants in all exercises noted the need to coordinate international responses to AI-bio threats, although approaches varied. While some groups of participants suggested publicly attributing threats to specific countries or companies, others advocated for more collaborative approaches focused on shared governance frameworks. Players in different exercises identified certain playbook components as important: requiring frameworks for information sharing with allies while protecting sensitive sources, diplomatic engagement for cooperation with China and other competitors, coordination mechanisms for multilateral responses to AI-bio incidents, and procedures for building international consensus on AI-bio governance.

Playbook 4: Attribution, Deterrence, and Response Escalation

Participants in eight exercises debated appropriate response thresholds throughout the exercise, particularly regarding the distinction between state and nonstate AI-bio threats and the escalation implications of various response options. The challenge of attributing AI-enabled biological attacks to specific actors or tools created uncertainty about the appropriate response in all exercises. As a result, participants emphasized the need for technical and legal standards for attributing biological attacks to actors using AI tools, escalation frameworks for responding to state versus nonstate AI-bio threats, and coordination between law enforcement and national security responses.

Playbook 5: Industry and Academic Engagement and Public-Private Partnership

In seven exercises, participants identified that private AI and biotechnology companies, as well as academia, would have a central role in both creating and mitigating AI-bio risks—and the U.S. government required new frameworks for government-industry coordination. Participants said that an effective response would require rapid mobilization of private-sector capabilities prepared ahead of an actual crisis—requiring managing potential conflicts between commercial interests and national security priorities. This necessitates frameworks for rapid engagement with AI and biotechnology companies during crises, information-sharing agreements with industry partners, procedures for compelling private-sector cooperation while respecting commercial interests, and mechanisms for leveraging private-sector capabilities for threat response.

Participants also discussed financial structures for supporting industry and identified government-underwritten research and development contracts as potential tools for supporting the creation of such countermeasures ahead of an AI-enabled pandemic.

Next Steps

These findings illuminate critical decision points and capability gaps that emerge when policymakers confront AI-enabled biosecurity crises. However, they represent only the beginning of a comprehensive research agenda designed to enhance U.S. preparedness for transformative AI developments.

We will expand this analysis through multiple parallel efforts. First, we will conduct additional iterations of the Viral Uplift scenario with enhanced design elements—including providing participants with more-sophisticated courses of action and predeveloped response playbooks for evaluation. These variations will test whether specific capabilities meaningfully improve crisis response effectiveness and identify which preparatory investments yield the highest returns.

Second, we will deploy our crisis simulation methodology across entirely different AI scenarios spanning economic disruption, military applications, and social stability challenges. This cross-scenario analysis will reveal whether the key issues and capability requirements identified in biosecurity contexts persist across other AI-related crises or represent domain-specific concerns. Such comparative analysis is essential for prioritizing government investments and avoiding narrow preparedness strategies that leave critical vulnerabilities unaddressed.

Appendix

Capabilities and Response Playbooks

Tables A.1 and A.2 summarize the capabilities and playbooks, respectively, that emerged from the exercises described in this report.

Table A.1. Summary of Proposed Capabilities

Capability	Description	Key Components
AI-bio risk assessments and technical evaluation	Enhanced capabilities to evaluate model performance in biological domains, assess technical barriers to biological weapon development, and monitor the evolving AI-bio threat landscape.	• Technical teams capable of evaluating AI model biological capabilities • Partnerships with AI and biotechnology companies for threat assessment • Standardized evaluation frameworks for AI-bio risk
AI model proliferation monitoring and analysis	Capabilities to monitor model development, distribution, and use while acknowledging direct control limitations	• Technical capabilities to analyze and characterize dangerous AI models • Monitoring systems for model downloads and use patterns • Honeypots and red-teaming to identify malicious use
Enhanced biological surveillance and early warning systems	Improved detection capabilities	• Pathogen-agnostic wastewater surveillance systems • AI-enabled genomic sequencing and analysis capabilities • Integration of biological and cyber threat monitoring
Rapid response biotechnology development for MCMs	Capabilities for accelerated development and deployment of MCMs	• AI-accelerated vaccine and therapeutic development platforms • Flexible manufacturing capabilities for rapid scale-up • Broad-spectrum countermeasures for unknown biological threats
International intelligence and cooperation framework	Enhanced capabilities for international cooperation on AI-bio threats	• Intelligence-sharing mechanisms with allies on AI-bio threats • Diplomatic engagement capabilities for AI-bio governance discussions • Monitoring of foreign AI-bio development programs and threat actors

Capability	Description	Key Components
Enhanced counterintelligence for AI-bio convergence	Specialized counterintelligence capabilities focused on AI-bio threats	• Monitoring of foreign intelligence activities targeting U.S. AI and biotechnology companies • Enhanced vetting and monitoring of those with access to dual-use AI-bio tools • Capabilities to detect and disrupt foreign acquisition of sensitive AI-bio technologies
Legal and regulatory response framework	Enhanced legal authorities and regulatory frameworks for AI-bio threats	• Updated legal frameworks addressing AI-enabled dual-use technologies • Regulatory frameworks for AI-bio research and development • Emergency authorities for rapid response to AI-bio incidents

Table A.2. Summary of Proposed Response Playbooks

Playbook	Purpose	Implementation
AI-bio crisis response protocol	Standardized procedures for responding to AI-enabled biological threats	• Rapid threat assessment and characterization procedures • Coordination mechanisms between health, security, and technology agencies • Communication protocols for public messaging and international engagement • Escalation criteria for different levels of response
Model containment and disruption operations	Procedures for model-focused responses while acknowledging control limitations	• Coordination with platforms and cloud providers for model removal • Information operations to disrupt threat actor access to dangerous models • Legal processes for compelling model removal or access restriction • Rapid identification and analysis of dangerous AI models
International cooperation and diplomatic engagement	Procedures for engaging with allies and competitors on AI-bio threats	• Diplomatic engagement protocols for cooperation with China and other competitors • Coordination mechanisms for multilateral responses to AI-bio incidents • Procedures for building international consensus on AI-bio governance • Framework for information sharing with allies while protecting sensitive sources
Attribution, deterrence, and response escalation	Procedures for attributing AI-bio incidents and calibrating responses	• Technical and legal standards for attributing biological attacks to AI tools • Escalation frameworks for responding to state versus nonstate AI-bio threats • Coordination between law enforcement and national security responses • Procedures for determining appropriate response levels

Playbook	Purpose	Implementation
Industry engagement and public-private partnership	Procedures for engaging with the private sector on AI-bio risks	• Framework for rapid engagement with AI and biotechnology companies during crises • Information-sharing agreements with industry partners • Procedures for compelling private-sector cooperation while respecting commercial interests • Mechanisms for leveraging private-sector capabilities for threat response

References

Halstead, John, "Managing Risks from AI-Enabled Biological Tools," *GovAI* blog, August 5, 2024.

Mitre, Jim, and Joel B. Predd, *Artificial General Intelligence's Five Hard National Security Problems*, RAND Corporation, PE-A3691-4, February 2025. As of August 5, 2025: https://www.rand.org/pubs/perspectives/PEA3691-4.html

Morris, Meredith Ringel, Jascha Sohl-Dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet, and Shane Legg, "Levels of AGI for Operationalizing Progress on the Path to AGI," *Proceedings of the 41st International Conference on Machine Learning*, PMLR 235, 2024.

Parson, Edward, "What Can You Learn from a Game?" in Ralph L. Keeney, Richard J. Zeckhauser, and James K. Sebenius, eds., *Wise Choices: Decisions, Games, and Negotiations*, Harvard Business School Press, 1996.

About the Authors

Anton Shenk is a quantitative research assistant at RAND. His research focuses on the economic and national security implications of artificial intelligence. He holds a B.S. in quantitative economics and mathematics.

Matt Chessen is a fellow at RAND's Technology and Security Policy Center. His research focuses on the geopolitical implications of transformative artificial intelligence. He holds a J.D. and an M.B.A. focused on Management Information Systems from the University of Arizona.

Barbara Del Castello is an associate physical scientist at RAND. Her research focuses on the impacts of technology maturity and diffusion on biological risk, as well as how life scientists are radicalized to violence. She holds a Ph.D. in genetics.

Gregory Smith is a policy analyst at RAND. His research focuses on emerging technologies, such as AI, and their implications for U.S. competitiveness and national security. He holds a J.D.

Richard S. Girven is a senior international defense researcher at RAND. His research focuses on emerging geopolitical risks, including the potentials and threats of artificial general intelligence. He holds a master of military arts and sciences in strategic studies.