NTI Paper

JANUARY 2026

A Framework for Managed Access to Biological AI Tools

SUMMARY

The convergence of artificial intelligence (AI) and the life sciences has driven the development of biological AI tools with a range of beneficial applications. However, some of these tools have also raised concerns among biosecurity experts that they could be misused by malicious actors to cause harm, including by making it easier to engineer dangerous pathogens. Managing access to these tools to ensure responsible use is critical for biosecurity not only because it decreases the possibility that they will be used for illegitimate purposes but also because managed access provides a foundation for oversight of how the tool is used. This report proposes a tiered, managed access framework and provides guidance on elements of the framework, including risk levels of biological AI tools, criteria for user legitimacy, and practices for verifying users.

Sarah R. Carter, PhD, and Greg Butchello

NTI:bio

Acknowledgments

The authors would like to acknowledge the support of our expert interviewees, who generously shared their time and expertise and whose contributions form the backbone of this report. We are also grateful to the experts who provided substantive feedback on the report, including Anthony Gitter, PhD; Cassidy Nelson, PhD; Phil Palmer, PhD; Ryan Ritterson, PhD; Dr. Toby Webster; and Jaime Yassif, PhD.

We would also like to acknowledge Nikki Teran, PhD, for her assistance with this project and Scott Nolan Smith on NTI's Communications team for managing the production of the report. Finally, the authors would like to thank Sentinel Bio for its generous financial support of this project.

Sarah R. Carter, PhD
Principal, Science Policy Consulting LLC

Greg Butchello
Program Officer, Global Biological Policy & Programs, NTI

Contents

Executive Summary

The convergence of artificial intelligence (AI) and the life sciences has driven the development of biological AI tools with a range of beneficial applications. However, biosecurity experts are concerned that some of these tools could be misused by malicious actors to cause harm, including making it easier to engineer dangerous pathogens. For example, tools used to guide vaccine development by providing insight into the characteristics of an emerging virus—such as how deadly it is or how readily it spreads among humans—could be misused to help design harmful variants of that virus. Managing access to these tools to ensure responsible use is critical for biosecurity not only because it decreases the possibility that they will be used for illegitimate purposes but also because managed access provides a foundation for oversight of how the tools are used.

This report builds on earlier NTI | bio work focused on biosecurity risks at the intersection of AI and biology and on potential guardrails to reduce risks related to biological AI tools.[1] It outlines a managed access framework that adopts two central principles: that access should be tiered on the basis of the level of risk associated with a biological AI tool and by the type of access that a user might be granted; and that the need for security should be balanced with the need for equitable access to the tool. One key idea that supports both principles is that providing users access to a biological AI tool through an application programming interface (API)—with appropriate oversight—can broaden accessibility while ensuring responsible use. The report also provides some guidance on core elements of the framework: providing tiered access to tools at different risk levels, establishing the risk level of each tool, and verifying legitimacy of the users of tools at different risk levels. However, these elements will need to be further developed, adapted by different communities of biological AI tool developers, and refined over time.

Recommendations

The following recommendations can be implemented by different parts of the life sciences community to support implementation of the managed access framework.

- **Funders of biological AI tools should**
  - » Offer low- or no-cost access to computational infrastructure to host models for developers who follow appropriate managed access procedures
  - » Fund the development of new managed access platforms, where needed
  - » Fund technical projects and practical workshops to support development of tools and best practices in support of managed access

- **Model developers, in partnership with biosecurity experts and others, should**
  - » Use a tiered risk framework to consider biosecurity risks during development
  - » Implement managed access approaches that are appropriate to the risk level of their tools
  - » Record lessons learned from implementing managed access and work with other model developers to identify best practices
- **Platforms that provide or host biological AI tools should implement managed access procedures consistent with this framework that**
  - » Support scientific innovation—for example, by enabling tools to be discovered, verified, used, adapted, compared with similar tools, and maintained over time
  - » Maintain equity and access for responsible users with transparent, defensible, and consistently applied criteria for users to establish legitimacy
  - » Expand access through secure, user-friendly APIs that include oversight to ensure responsible use

As biological AI tools continue to advance, adopting this framework can reduce risks, improve transparency, support informed decisions about managed access, and promote beneficial use. Platforms that support managed access can provide many benefits for tool developers and users by ensuring that useful tools can be discovered, accessed, compared with other tools, and maintained over time. This approach also can provide assurances to funders, policymakers, and the broader public that the life sciences community is pursuing its scientific goals responsibly and with due diligence for safety and security.

Introduction

The convergence of AI and the life sciences is driving the development of biological AI tools that support fundamental research as well as applications spanning therapeutics, agriculture, and biomanufacturing—offering significant benefits for society.[2] However, biosecurity experts have raised concerns about some of these tools because the tools could be misused by malicious actors to cause harm, including making it easier to engineer dangerous pathogens.[3] This report outlines a tiered, managed access framework for biological AI tools that aims to reduce these biosecurity risks while ensuring that the benefits can be realized.

Many developers in the life sciences community have committed to addressing risks related to biological AI tools.[4] However, to date, there have been few resources developed to help developers determine which tools may contribute to biosecurity risks and what to do when risks are identified. This report aims to address this gap. One critical approach to reduce risk is managed access—allowing only validated users to access tools that contribute to biosecurity risks.[5] Most directly, this approach decreases the probability that tools will be used for illegitimate purposes. Managed access is also foundational to other risk-reduction strategies, including monitoring and the use of built-in technical guardrails.[6] If a tool is released fully and openly, then these strategies can be avoided or removed.

This report outlines a managed access framework for responsible use and dissemination of biological AI tools that has a tiered access approach at its core. Importantly, biological AI tools—that is, AI-enabled tools trained on biological data and intended to provide predictions, insights, or designs related to biology—are diverse, and many do not significantly contribute to biosecurity risks. To balance the need for security with the need for access by legitimate users, the framework is tiered by the level of risk associated with a biological AI tool and by the level of access that a user might be granted. This report also provides some initial guidance on the elements of the framework, including how to establish risk levels for tools, criteria for legitimate users to access tools at different levels, and practices for verifying those users. However, this guidance is not prescriptive because the implementation of each element of the framework is likely to change over time, even as the framework remains the same.

Managed Access in the Life Sciences

Managed access to in silico resources and tools has been implemented in the life sciences in many different contexts and for various reasons. For example, many policies govern access to human genome data: users of that data are generally required to meet stringent criteria for access,[7] and experts have recommended that access to machine learning models trained on this sensitive data should also be managed.[8] Some pathogen genome databases also have access controls; for example, the Global Initiative on Sharing All Influenza Data (GISAID) requires registration with an institutional email address and agreement to its terms of use before data can be accessed.[9] Access to many academic resources and tools are restricted to individuals affiliated with the institution that provides such resources. The framework described here draws on these existing practices and incorporates lessons learned from a series of managed access case studies in the life sciences.

As another example, commercial developers of biological AI tools often incorporate managed access practices. Many companies, including those developing pharmaceuticals and therapeutics, consider their data and related tools to be proprietary, and they keep them carefully guarded. Companies that want their tools to be broadly available often adopt tiered, managed access approaches in which users must meet more stringent criteria to gain more complete access to the tool. Through an application programming interface, Google DeepMind provides access to a version of AlphaFold3, its protein structure prediction model, to any user,[10] but users who demonstrate an academic or other noncommercial affiliation can access the model more fully by downloading its weights.[11] Similarly, Evolutionary Scale offers its ESM3 biological foundation model through an API but openly provides a smaller version (ESM3-open) for academic research purposes.[12] In these commercial contexts, decision-making about managed access often includes considerations about responsible development, security, and the potential for misuse of the tools.[13]

Managed access to biological resources to reduce potential biosecurity risks has also become a focal point in academic publishing.[14] For example, Microsoft, in collaboration with the International Biosecurity and Biosafety Initiative for Science (IBBIS), recently announced a tiered access approach[15] for a set of biological data and tools that were developed for a study published in *Science*.[16] The study tested the vulnerabilities of nucleic acid synthesis screening procedures to obfuscation of sequences by biological AI tools, and the study team determined that some tools and resources developed for the study pose dual-use risks because they could be used to design sequences to bypass or probe these biosecurity practices.

The adoption of the managed access framework will help the community reduce risks, improve transparency, support informed decisions on managed access, and promote beneficial use.

The tiered, managed access framework described in this report is consistent with these existing tiered approaches. This managed access framework for biological AI tools is intended to be durable and adaptable to different contexts and can guide developers, funders, publishers, and others in the life sciences community to a shared understanding of responsible use. By separating out questions of tiered access, risk levels based on misuse-relevant capabilities of different types of tools, and criteria for legitimate users and legitimate use, this report aims to provide structure to what has been a wide-ranging and sometimes difficult discussion within the life sciences community. The report also highlights areas where additional support is needed to develop resources and best practices for decision-making and implementation and it includes considerations about platforms for managed access. Recommendations for the life sciences community, including funders, are included in the final section of the report.

As biological AI tools continue to advance, the adoption of the managed access framework will help the community reduce risks, improve transparency, support informed decisions on managed access, and promote beneficial use. Platforms that support managed access will ensure that useful tools are discovered, accessed, compared with other tools, and maintained over time. This approach also will provide assurances to funders, policymakers, and the broader public that the life sciences community is pursuing its scientific goals responsibly and with due diligence for safety and security.

Project Methodology

The development of this managed access framework draws on previous work as well as multiple rounds of interviews and feedback from a range of experts in biosecurity, AI, biological AI model development, and the broader biosciences. An initial phase of this project focused on case studies for managed access. In this phase, six semi-structured interviews were conducted with developers and funders of biological AI tools, databases, and other resources who had implemented or considered implementing managed access procedures. Lessons learned from these case studies informed the initial development of the managed access framework. (Appendix A discusses lessons learned, and Appendix B describes each case.)

A second phase of the project included semi-structured interviews with 15 individuals who provided their insights on managed access and input on an initial framework document. A full draft report was circulated before a virtual workshop on September 24, 2025, during which feedback was solicited on each part of the framework. A list of project participants, including interviewees, workshop attendees, and others who provided substantive feedback on the framework, is provided in Appendix C. Although this report was developed with input from those experts and incorporates diverse perspectives, it was written by its named authors alone.

Proposing a New Managed Access Framework

The proposed managed access framework for biological AI tools supports biosecurity by providing a tiered approach for access to tools at different levels of risk. By verifying that users meet specified criteria before accessing tools at higher levels of risk, developers and others can reduce the possibility that the tools are misused to cause harm. At the same time, ensuring that the benefits of tools, at all levels of risk, can be fully realized is important. To maximize benefits while reducing risks, the managed access framework described here adopts two central principles:

- **Access should be tiered on the basis of the tool's risk level.** Biological AI tools at a low risk level should have few or no access restrictions, while tools at a higher risk level should have more stringent requirements for access. Providing access to tools at different levels—for example, by providing access through an API rather than providing the full source code, data, and model weights—creates opportunities for oversight to reduce risks while enabling access to a broader range of users.
- **The need for security should be balanced with the need to provide equitable access.** The benefits of these tools, including advances in biosecurity and public health applications, depend on the ability of scientists and other users to access them. Equity requires that criteria for establishing user legitimacy or legitimate use of these tools be carefully considered so that any researchers or others with legitimate uses, even under diverse circumstances, can access the tools.

These principles and the elements described in the following sections constitute a durable managed access framework. The sections on tiered access, risk levels, and legitimacy verification provide descriptions and guidance on how each element could be implemented. Given the diversity of biological AI tools and the contexts in which they are developed and used, it is likely that the guidance provided in each of the sections will need to be adapted to best meet the needs of different developers. The framework also should be updated over time as technology advances, uncertainties about risk are resolved, and best practices for implementation are established for each of the following three elements.

Tiered Access to Biological AI Tools

Biosecurity risks related to biological AI tools depend not only on the capabilities of the tool, but also on the level of access to the tool that a user is granted. Among many model developers and others in the life sciences, a model is fully "open source" if the model's code, data, and weights are all publicly available. These model components contribute to decision-making for managed access in different ways:

- **Code:** The code is the foundation of the model, including its architecture and algorithms, and it captures the developers' innovations and advances in modeling techniques. Access to the code alone may be adequate for evaluations of a developer's contributions to the field, for example, for publication purposes. The code needs to be combined with training data to produce a trained model.

- **Training data:** The training data are central for biological AI tools. If a model's code plus training data are open source, then a technically competent user with adequate compute can create a model of equivalent performance. Still, releasing only the code plus training data can be an effective way to manage access to a model if the computational resources required to train it are prohibitive. If a developer has trained their model using data that are not publicly available, then establishing managed access for the data also can serve as a means for managing access to the model.
- **Model weights:** The model weights alongside the code represent the fully trained model. Any experimental validation, predictions or demonstrations of generalizability, and other features would be captured with the model weights. Therefore, access to the model weights provides the most unhindered access to the full capabilities of the model and provides a critical opportunity for managing access.

Access to a fully trained model, including all code, data, and weights, enables a technically capable user with access to sufficient computing infrastructure to implement the model, further develop or fine-tune the model, and share it with others without oversight. If built-in guardrails have been incorporated into the model—such as watermarks,[17] screening systems, methods to capture metadata associated with the model,[18] or others[19]—they can be removed. When all components of a model are shared openly, then it is likely that some version of it will always be openly available—that is, the developer and others will be unable to retract or retrieve it in the future.

The core of the tiered access framework is provided in Table 1, which indicates what types of users should be granted access to biological AI tools at different risk levels. Because access to the full model provides full capabilities with no opportunity for oversight, only biological AI tools at risk level 1 (very low level of misuse-relevant capabilities) should be fully and openly released to all users. For tools at risk levels 2, 3, and 4, users should meet some criteria for legitimacy. The following sections provide guidance on what types of tools might fall into each risk level, as well as practices for how user legitimacy at each level can be verified.

A developer can reduce the risk of misuse of their biological AI tool by choosing to provide access through a hosted API. An API allows a user to interface with a biological AI tool without gaining full access to the underlying model. This option allows the developer to maintain control of the model, to monitor its use, and to ensure that any built-in guardrails are maintained. A few interviewees noted that this arrangement, when combined with basic checks of user legitimacy, can reduce the possibility of frontier large language models (LLMs) or other "bots" accessing the tool. As shown in Table 1, developers who provide access through an API can use less stringent criteria for user legitimacy—that is, they can provide access consistent with a lower risk level for the model—compared with those developers who provide access to the full model.

Table 1. Tiered Access Framework

	Provide Full Model (Code, Data, and Weights), with Appropriate Limits on Further Sharing	Provide Access through an API, with Appropriate Oversight
Risk Level 1: Very Low Level	All users	All users
Risk Level 2: Low Level	Identifiable, legitimate users	All users
Risk Level 3: Medium Level	Identifiable, legitimate users with a reason to access the tool	Identifiable, legitimate users
Risk Level 4: High Level	Identifiable, legitimate users with a specific project that requires access to the tool	Identifiable, legitimate users with a reason to access the tool

Model access through an API can expand the number of users with access to a biological AI tool and is an important means of achieving truly equitable access. Open access to the components of a tool does not enable all users to successfully elicit meaningful outputs from the model or build on the model's capabilities. In addition to computational resources to run the model, some level of technical expertise is required to understand and implement the code and to interface with the trained model. In this way, the "openness" of a model is not the same as "accessibility" to users.[20] Managed access through an API can expand access to a broader range of people and level the playing field by improving the usability of a tool, decreasing the technical expertise required, and providing computational resources to run the model. Several interviewees pointed to the Galaxy Hub for bioinformatic resources[21] as an example of expanded access to tools through APIs. APIs can also be designed to meet the needs of users who aim to further develop the tool. Although APIs are typically implemented in a way that prevents access to the underlying model weights, some enable the user to provide additional data to fine-tune the model for specific purposes.

In addition to expanding access to a biological AI tool, APIs can reduce the risk of misuse because they provide an opportunity for the host to oversee the tool and how it is used. The type of oversight that is appropriate for a tool or workflow will vary based on the type of tool, its risk level, its number of users, and the context in which it is used. Several interviewees pointed out that even simple monitoring—for example, of the number or pattern of queries from a user—can provide a way to flag unusual use. High-volume users could be subject to additional verification. Oversight can also include the use of guardrails, such as screening of user requests or inputs to the tool for features that may raise biosecurity concerns or that are outside the anticipated or typical use. For example, a flag might be raised when a user of a general-purpose tool requests designs related to a high-risk pathogen. More complex types of monitoring or oversight could be implemented, including the use of biosecurity "agents" to detect other patterns of use that may pose biosafety or biosecurity risks. If an API allows fine-tuning of the tool, using guardrails and monitoring can help ensure that this process does not introduce additional biosafety or biosecurity risks. Any flags that are raised during oversight can be logged and can form the basis for follow-up with the user or refusal of the model to generate outputs. A few project participants raised the possibility that monitoring could be linked

with law enforcement, when warranted. Many of these approaches are underdeveloped and will require focused investment to ensure that biosecurity monitoring and oversight capture meaningful risks without hindering beneficial uses.

When providing access to a fully trained model, the developer should consider the conditions under which the tool could be further shared, which will vary depending on the risk level of the tool. Licensing agreements can set expectations for how a model should be used and disseminated. For tools at risk level 1, an open-source licensing agreement may be appropriate—for example, the MIT license, which "grants permission to use, modify, and distribute the software, with the condition that the original copyright notice and the license text are retained in the redistributed software."[22] For tools at risk level 2, a licensing agreement should ensure that the tool is shared only with those who are identifiable, legitimate users (as described in the following section Verifying Legitimacy for Users of Biological AI Tools). At higher risk levels (3 or 4), a licensing agreement should require that tools not be further shared. Particularly for those tools, developers could pursue additional methods to ensure that the tools are not inappropriately shared (such as a requirement for a signature rather than a simple "click-through" or "clickwrap" agreement to terms) or they could pursue technical approaches that enable access to model weights in a hosted, protected infrastructure that prevents downloading or exporting of the tool. For any tool that incorporates built-in safeguards, such as screening systems, watermarks, or approaches for metadata collection, developers should consider adding clauses to licensing agreements or pursuing technical approaches to better ensure that safeguards are maintained. As mentioned in the lessons learned from the case studies (Appendix A), templates or standardized licensing agreements for access to biological AI tools do not exist but would be helpful for the community.

Many biological AI tools provide substantial potential benefits, including for pandemic preparedness and response, and it is important that those tools are accessible to legitimate users. Access, often including access to the full code, training data, and weights, is important not only to ensure that the tools can be used as intended, but also to support the scientific imperatives of transparency and reproducibility, peer review, and further refinement and development. The section Verifying Legitimacy for Users of Biological AI Tools includes examples of practices that a developer or platform can adopt to verify user legitimacy and legitimate reasons to use tools at different risk levels. In making those determinations, the developer or platform should accommodate these scientific imperatives as legitimate reasons for access whenever possible.

Risk Levels for Biological AI Tools

A central concern for biological AI tools and integrated workflows that include biological AI tools is that they may be misused by a malicious actor to design pathogens, including variants of high-risk pathogens not found in nature.[23] The Centre for Long-Term Resilience (CLTR) and RAND Europe recently published a detailed rubric for assessing "misuse-relevant capabilities" for biological AI tools that could contribute to this concern.[24] Although best practices for evaluating misuse-relevant capabilities of biological AI tools and assigning risk levels are not established, this rubric serves as a helpful starting point. Previously, biosecurity experts outlined in broad terms how these tools might contribute to risk,[25] capabilities that may be dual-use,[26] and considerations for the types of data that, when generated or included in the training data for a model, may increase biosecurity concerns.[27] Well-resourced tool developers in industry, nonprofit,

and government labs have conducted their own assessments of biological AI tool capabilities, but those assessments have been ad hoc and highly specific to their own tools.[28]

Risk assessment of biological AI tools and workflows can also draw on criteria and resources developed in the context of synthetic biology, dual-use research of concern (DURC), and related frameworks that identify experimental research that may pose additional risk and should be subject to additional oversight. The National Academies report, *Biodefense in the Age of Synthetic Biology,* outlines a framework for understanding features of a new capability that may increase concerns about its misuse.[29] Additional resources include the seven experiments of concern from the National Academies report *Biotechnology Research in an Age of Terrorism* (also known as the Fink report),[30] guidance from the U.S. government in 2024 on DURC and enhanced potential pandemic pathogens,[31] Stanford University's Visibility Initiative for Responsible Science,[32] and a checklist developed by the Netherlands Biosecurity Office.[33]

Model access through an API can expand the number of users with access to a biological AI tool and is an important means of achieving truly equitable access.

Importantly, the risk of misuse of a biological AI tool depends both on the misuse-relevant capabilities of the tool as well as the likelihood that an actor would successfully use it to cause harm. A high-quality risk assessment of a biological AI tool should incorporate a wide range of contextual expertise, such as knowledge of bioweapons development processes and bottlenecks, resources that may be used alongside the tool being assessed, and types of adversarial actors, including their likely capabilities and possible motivations. Ideally, the assessment would be conducted in collaboration with biosecurity experts and others who can offer this expertise. Although this type of assessment most accurately establishes a level of concern for a tool, many developers and others in the community do not have access to this expertise nor the resources necessary to support this type of assessment.

In the absence of a full risk assessment, the guidance provided here outlines potential misuse scenarios (see Table 2), risk factors, and examples (see Box 1) to help developers determine an appropriate risk level for their tools and a corresponding managed access approach. A central question is to what extent does the model have misuse-relevant capabilities? Table 2 lists potential misuse scenarios for eight types of biological AI tools. By carefully considering the extent to which their tools could be misused to realize one or more of these scenarios, developers can determine whether their tool has misuse-relevant capabilities and therefore should be considered for a higher risk level. The CLTR-RAND Europe report provides additional definitions and guidance to determine a tool's misuse-relevant capabilities.[34]

Table 2. Potential Misuse Scenarios

Tool Category	Potential Misuse Scenarios
Viral Vector Design	Designing proteins that enhance misuse-relevant properties for viruses with significant weaponization or harm potential
	Designing biological agents with increased stability under misuse-relevant conditions for formulation, storage, and delivery as a weapon, for agents with significant weaponization or harm potential
Protein Engineering	Engineering proteins that enhance misuse-relevant properties for agents with significant weaponization or harm potential
	Engineering more harmful protein toxins by altering known toxins or designing novel ones
	Engineering proteins such as surface proteins with increased stability under misuse-relevant conditions for formulation, storage, and delivery as a weapon, for agents with significant weaponization or harm potential
Small Biomolecule Design	Identifying new harmful, small molecule toxins by altering known toxins or designing novel ones
	Designing small molecules that enhance pathogen effects or delivery (referred to as enhancer molecules)
Genetic Modification and Genome Design	Designing pathogen genomes to create agents with enhanced misuse-relevant properties
	Designing pathogen genomes to create agents with enhanced stability under misuse-relevant conditions
	Enabling evasion of pathogen detecting and screening methods, including those used for biosurveillance and synthetic nucleic acid screening
Pathogen Property Prediction	Identifying pathogen characteristics that could enhance virulence
	Identifying features that alter immune modulation or antimicrobial resistance
	Identifying altered tropism and zoonotic spillover potential
Host–Pathogen Interaction	Predicting host–pathogen interactions that enhance misuse-relevant properties of pathogens with significant weaponization or harm potential
Immune System Modeling and Vaccine Design	Identifying genetic components that enhance misuse-relevant properties of a pathogen with significant weaponization or harm potential
	Identifying strategies to increase the stability of pathogens in misuse-relevant conditions
	Identifying novel immune modulation pathways or targets
Experimental Design, Simulation, and Automation	Designing and executing workflows to optimize production efficiency, yield, and scale of biological agents with significant weaponization or harm potential
	Autonomously analyzing, characterizing, and validating production outcomes without human expertise

Source: Adapted from Toby Webster et al., *Global Risk Index for AI-Enabled Biological Tools* (Centre for Long-Term Resilience and RAND Europe, September 9, 2025), https://doi.org/10.71172/wjyw-6dyc.

The appropriate risk level of a tool will vary based on the extent of its misuse-relevant capability and whether the tool also possesses one or more additional risk factors, including the following:

- The model uses new, experimentally generated data rather than publicly available data. Data are a key bottleneck for biological AI tools, particularly for pathogen-related and misuse-relevant applications.[35] Therefore, tools that are trained or validated using new, experimentally generated data may have unique capabilities that could be misused. Models that incorporate new, pathogen-related data generated at larger scales or in multiple cycles linked to model development and refinement might warrant a high risk level.
- The model can produce novel or previously undocumented designs. A model that can provide functional designs for molecules that are highly divergent from those found in nature might be considered for a higher risk level. Models that have been experimentally validated for such designs may warrant a higher risk level.
- The model is generalizable. A model that can provide designs or predictions across a wide range of biological organisms or their subcomponents may warrant a higher risk level.

Box 1. Risk Levels and Examples

The risk level of a biological AI tool describes the extent to which it has misuse-relevant capabilities and incorporates additional risk factors.

RISK LEVEL 4

High Level of Misuse-Relevant Capabilities

Example: A model that is trained on high-risk pathogen genomes validated with new, experimentally generated data on functional outcomes of different variants.

RISK LEVEL 3

Medium Level of Misuse-Relevant Capabilities

Example: A model that is trained on genomes of an adeno-associated virus (a commonly used vector for gene therapy) that designs variants that can evade existing immunity. It is unlikely to generalize to other viruses.

RISK LEVEL 2

Low Level of Misuse-Relevant Capabilities

Example: A biological foundation model that is trained on large amounts of diverse protein or genomic sequences and is highly generalizable across many domains of life but has not been experimentally validated.

RISK LEVEL 1

Very Low Level of Misuse-Relevant Capabilities

Example: A model that is trained on protein sequences and structural data and provides predictions or scores that indicate whether a protein will interact with other proteins or small molecules.

This approach to assessing risk and assigning risk levels for biological AI tools is intended to provide a starting point and should be revisited and refined over time. As developers and the broader life sciences community gain experience with this type of analysis and decision-making, they will be able to develop additional criteria and heuristics. Also, as technology advances and the landscape of risks changes over time, biosecurity experts and others should continue to update guidance on misuse scenarios, misuse-relevant capabilities, risk factors, and other considerations for understanding risks related to these tools.

Understanding risk factors and their implications for managed access can help drive decision-making about the design and dissemination of a tool before it is trained. Biosecurity experts have highlighted the need for evaluation of risks and model capabilities early in the development cycle of biological AI tools.[36] In addition to careful consideration of the tool's purpose, design, training data, and validation, those developing biological AI tools at higher risk levels (3 or 4) can integrate dissemination plans and development of APIs into their projects, if needed. Importantly, the number of biological AI tools at these higher risk levels is likely to be lower than those at lower risk levels.[37]

Verifying Legitimacy for Users of Biological AI Tools

Table 3 outlines practices that could be implemented to establish legitimacy for access to biological AI tools at each risk level, as described in the tiered access framework (see Table 1). Methods to verify institutional affiliation, legitimacy as part of the scientific community, and legitimate use have been explored extensively in the context of customer screening by nucleic acid synthesis providers,[38] and resources have been developed in that context to help with decision-making.[39]

The practices described in the following paragraphs draw on this existing biosecurity framework and take an analogous approach. Because biological AI tools are strictly in silico and do not provide reagents to physically instantiate their designs, tools in risk level 1 do not require any measure of legitimacy for use or further distribution. For risk level 2, practices include checking for an institutional affiliation, which is similar to best practices adopted by nucleic acid synthesis providers for customers who order benign nucleic acid sequences (i.e., sequences that show no homology to pathogen or toxin sequences) as well as practices in many other life sciences contexts (e.g., GISAID, Addgene, ThermoFisher, Galaxy Hub's Jupyter Notebooks). Practices for risk level 3 include additional checks of a user's identity and verification that the user has some link to the scientific community beyond a simple affiliation, while risk level 4 includes verification of project plans. These practices draw on methods that nucleic acid providers use when a customer orders a nucleic acid sequence that shows homology to pathogen or toxin sequences. Notably, practices by nucleic acid providers for these high-risk orders go beyond the practices outlined here, even at the highest risk level, and often include requests for documentation of institutional legitimacy and signed approvals from institutional biosafety officers.

Table 3 assumes that access to a biological AI tool will include access only to in silico resources. Some biological AI tools may be incorporated into experimental, "wet-bench" capabilities—for example, laboratory robotics or lab-in-the-loop operations—that may generate additional risk. Access to these types of tools may require additional oversight to ensure that they are not accidentally or deliberately misused.

Table 3. Practices to Verify User Legitimacy for Access to Biological AI Tools at Each Risk Level

	Risk Level			
	1	2	3	4
All Users: No Checks Necessary	✓			
Identifiable, Legitimate Users • Automated check of email address or phone number • Check for affiliation with an institution or group with a legitimate interest in biological AI tools. Items to verify include the following: - Institutional email address - Membership or participation in a relevant group - Reasonable explanation of affiliation - If unfamiliar with the institution or group, a website or other publicly available resources		✓	✓	✓
Identifiable, Legitimate Users with a Reason to Access the Tool • Verify the user's identity to a higher degree of certainty. Options include the following: - Contact the user via phone or video conference - Check government-issued identification • Check that the user is part of the scientific community. Options include the following: - Check publications, ORCID, or participation in conferences - Check publicly available information to determine that the user's company or institution has a relevant mission - Check for a reasonable explanation if other criteria do not provide assurance • Check for completion of training module on safeguarding biological AI tools against misuse (if available).			✓	✓
Identifiable, Legitimate Users with a Specific Project That Requires Access to the Tool • Check documentation and verify project plans for legitimate use of the tool. Items to check include the following: - Grant or other funding documentation - Institutional support or oversight of the project - Contact another individual at the institution with knowledge of project plans				✓

A central concern, both for nucleic acid synthesis providers and among interviewees for this project, is ensuring equitable access. Measures of "legitimacy" are often biased toward users at well-resourced, established institutions and against users from nontraditional, newer, and lower-resourced organizations, particularly in the Global South. Multiple interviewees emphasized the need for flexibility in establishing criteria to better account for legitimate users in these broader contexts. One interviewee stated that it was important to have a human in the loop rather than relying on automated checks to evaluate the information provided by users to make determinations about legitimacy. Another project participant highlighted the importance of ensuring free or low-cost access for all legitimate users.

In Table 3, the examples of practices provided in each section are meant to provide guidance but should not be considered prescriptive. As is the case for customer screening by nucleic acid providers and in other contexts, these determinations require due diligence but will ultimately depend on the good-faith judgments of those making the decisions.

Implementation of these practices for tools at risk level 2 could involve a simple registration process that includes confirmation of an email address and a check for institutional affiliation. Many interviewees believe that this process could reduce risks without adding a significant burden for users and that it is similar to other familiar processes. However, implementing and consistently maintaining this type of approach over time could be difficult for some biological AI tool developers, particularly those at institutions with fewer resources or high levels of staff turnover. To verify legitimacy criteria at higher risk levels (3 or 4), developers would require additional resources and ongoing support for these more complex determinations. Although some tool developers are willing and capable of managing access at these levels, establishment of managed access platforms for biological AI tools would increase the efficiency, clarity, and consistency of this approach to risk reduction.

Considerations for Platforms

Here, a "platform" for biological AI tools refers both to the computational infrastructure that houses one or more tools as well as the entity that takes responsibility for establishing governance and managing access to those tools. A platform could simply house downloadable code and other components of a biological AI tool and it could incorporate a tiered access approach by providing access to users on the basis of some criteria, as described previously. Rules for access could differ based on what is made available for download; for example, the code, training data, and weights could each have different criteria. Alternatively, a platform could host a biological AI tool for external users through an API, as described in a previous section. Hosting a tool or a suite of tools can expand access to the tool while reducing risk by providing opportunities for oversight to ensure responsible use.

A platform could support tool developers by helping to establish appropriate governance and decision-making for their tools, including cybersecurity requirements, licensing agreements, practices for monitoring of tool access or use, and consistent verification of user legitimacy and legitimate use of tools at different risk levels. Some interviewees emphasized that a platform could best meet the needs of developers by ensuring that the scientific needs for peer review and transparency are met—for example, by providing full access to a model's source code, data, and weights when needed. Ideally, the platform could provide assurances to the developer that its managed access approach is consistent with responsible development principles, data sharing requirements from publishers and funders, and broader community expectations.

Some types of platforms that distribute or host biological AI tools may be able to draw on existing networks and communities to more easily establish a user's legitimacy and legitimate use of the tool. For example, a platform that houses a suite of tools could be established for a community of users based on a common funder, a collaborative purpose or goal, or shared interest in tool development. The case studies described in Appendix B largely follow that approach. The criteria for legitimacy described in Table 3—including an "affiliation with a legitimate institution or group" and a "legitimate reason to use the tool"—could then be verified or supported on the basis of membership, collaboration, or sponsorship within that community. That approach may be particularly well suited for tools at higher risk levels for which information about specific users and projects informs decision-making about access. The Coalition for Epidemic Preparedness Innovations (CEPI) has already committed to a broad, equitable access platform for its biological AI tools,[40] and other collaborative research organizations and funders could follow. One project participant suggested that journals should create or otherwise support managed access platforms for the tools and data that they publish.

A new, independent platform also could be established to house biological AI tools and incorporate a managed access approach. By providing a range of features and incentives, this type of platform could meet the needs of both users and developers. Many interviewees believe that a central feature driving adoption would be hosting biological AI models through an API and simplification of the interaction with the models. The platform could also provide tutorials, guidance, and assurances that tools are up to date, or it could enable the community of users to contribute resources. Several interviewees raised the possibility that a platform could host a range of biological AI tools so that multiple similar tools, including smaller or more niche tools, could be easily compared and adapted for different tasks. The use or development of

performance benchmarks would facilitate this type of comparison and enable users to more easily determine the state of the art for these tasks. Platforms could also enable further development of tools—for example, by providing structured data and other support for model training.[41] Developers of biological AI tools could also benefit from many of the features of a managed access platform, particularly if the platform enables them to track use of their tool over time. That type of platform may be most relevant for biological AI tools at lower risk levels, for which there are likely to be many similar tools and many types of users.

While platforms that host biological AI tools offer valuable opportunities, successfully deploying them would require addressing some challenges. Interviewees noted that establishment and maintenance would be expensive. Costs would include computing expenses, which could be substantial if the platform is designed to host many types of tools. Additional costs would come from the technical development of APIs and other resources. There would also be ongoing expenses to maintain software and to meet the needs of a community of users and developers. Some interviewees also raised broader issues related to trust. For example, if a U.S. national laboratory or other U.S. government entity established a platform, then it could be difficult for some international users and developers to participate. One interviewee pointed out that some scientists may be hesitant to provide identifying information in some cases out of fear that their research interests may make them a target of funding cuts or intimidation.

Platforms for biological AI tools already exist, including many that integrate some type of managed access. Interviewees pointed to several examples: the hosting platform Hugging Face provides access to a wide variety of AI models (not just those related to biology) and can be configured to require some registration and licensing agreement for access.[42] A variety of platforms specifically for biological AI tools have been established for commercial purposes by companies that provide the tools with user-friendly interfaces or combined with some additional service. Examples include Neurosnap,[43] Rowan,[44] Tamarind Bio,[45] Levitate Bio,[46] and others.[47] Companies such as TeselaGen[48] and Benchling[49] integrate many types of tools alongside more comprehensive support for laboratory-based research and development. Although these platforms were not developed for the purpose of reducing biosecurity risks, their managed access practices may meet some of the same needs.

To effectively reduce risks related to biological AI tools, it will be important to leverage existing platforms and communities. These platforms and communities are diverse, and elements of the managed access framework described in this report are likely to be incorporated in a variety of ways. Engagement between biosecurity experts and tool developers will be critical to ensure that tools that may contribute to biosecurity risks are identified and that appropriate managed access approaches are implemented in a way that meets the needs of each community.

Recommendations

This report offers a framework for managed access to biological AI tools while recognizing that, to be successful in a rapidly changing environment, each element of the framework should be revisited and refined over time. The risk considerations and methods for assigning tools to different risk levels should evolve as the biosecurity community gains a better understanding of how different types of tools might be misused to cause harm.

Best practices for granting users tiered levels of access and decision-making about user legitimacy and legitimate use of tools also should be developed and adapted to meet the needs of different communities of biological AI tool developers and their users. This process should be conducted in collaboration with publishers, funders, and others in the life sciences community to ensure that scientific needs such as peer review and transparency can be met, even for tools at the highest risk level.

Best practices... should be developed and adapted to meet the needs of different communities of biological AI tool developers and their users.

The recommendations here outline steps that different members of the life sciences community can take to support this framework.

- ## Funders should do the following:
  - » Offer low- or no-cost access to computational infrastructure to host models for developers who follow appropriate managed access procedures.
  - » Fund the development of new managed access platforms, where needed.
  - » Fund technical projects and practical workshops to support development of tools and best practices in support of managed access. These include standardized licensing agreements or templates for tools at different risk levels; effective approaches for oversight of biological AI tools through an API; resources for risk assessment and assignment of risk levels, including publicly available rubrics vetted by biosecurity experts; and decision-support tools for verifying user legitimacy, institutional affiliation, and legitimate use of biological AI tools in different contexts.
- ## Model developers, in partnership with biosecurity experts and others, should do the following:
  - » Use a tiered risk framework to consider biosecurity risks during development.
  - » Implement managed access approaches that are appropriate to the risk level of their tools.
  - » Record lessons learned from implementing managed access and work with other model developers to develop best practices.

- Platforms that provide or host biological AI tools should implement managed access procedures consistent with this framework that also include the following:
  - » Support scientific innovation, for example, by enabling tools to be discovered, verified, used, adapted, compared with similar tools, and maintained over time.
  - » Maintain equity and access for responsible users with transparent, defensible, and consistently applied criteria for users to establish legitimacy.
  - » Expand access through secure, user-friendly APIs that include oversight to ensure responsible use.

Managed access will be critical for reducing biosecurity risks related to misuse of biological AI tools. By working to develop best practices for each element of this framework—tiered access, risk levels, and practices to verify legitimacy—developers of biological AI tools and the broader life sciences community can reduce risks while maintaining the benefits of these tools. Over time, as implementation of this framework becomes more commonplace, as platforms develop, and as tools and resources are generated to support best practices, funders, governments, and other stakeholders should establish incentives for broader adoption.

Appendix A. Case Studies: Key Challenges and Lessons Learned

To inform the development of a managed access framework for biological AI tools, it is important to learn from existing approaches. Managed access case studies were explored through semi-structured interviews with individuals who have implemented or are actively considering managed access approaches for biological resources, AI models, and related tools. Interviews included questions about how managed access approaches are or could be implemented, how well these approaches meet user and developer needs, how equitable access is or could be maintained, and any challenges encountered or lessons learned. More information on each of the six case studies is included in Appendix B. Key challenges and lessons learned are summarized here.

Resources for Managed Access

Resources are required to establish and maintain an appropriate managed access approach. Among the case studies, smaller organizations pointed to costs such as hosting fees (i.e., cloud computing) and legal fees for licensing agreements. Implementing the managed access approach would require personnel to establish the system and verify user legitimacy on an ongoing basis. Previous studies have highlighted this challenge,[50] particularly for academic tool developers who do not have funding for managed access and do not work in institutions with the infrastructure, personnel, or expertise for implementation. Larger organizations are better equipped to implement a managed access system. For example, case study interviewees at U.S. national laboratories were confident that they had the computing infrastructure to host even very large biological AI models; legal, contracting, and oversight support to establish a managed access approach; and personnel and procedures for verifying collaborators and users of tools. Although national laboratories and other U.S.-based organizations may be well equipped to operate managed access platforms, one case study interviewee pointed out that such platforms may not work for biological AI models, databases, and resources that are intended to be shared broadly with international users. Platforms for these users may need to navigate additional complexities on governance and establishing trust in an international context.

Biosecurity Risk Assessment

Case study interviewees believe that their resources or tools could pose some level of biosecurity risk, but in many cases, this determination was made on an ad hoc basis and with limited guidance. Additional tools and resources are needed for systematic risk assessment, how to balance risk with the benefits of open sharing, how to define discrete subsets of data that may warrant control due to biosecurity risks, and the implications of different levels of risk for managed access. Often, biosecurity (i.e., the risk that the data, biological AI tools, or other resources might be misused to cause harm) is only one consideration among many that drive decision-making about whether and how to implement a managed access approach. Commercial considerations, licensing agreements with developers of databases or tools, and policies or norms related to data sharing can also play a role.

Verification of User Legitimacy

Some case study interviewees highlighted the challenge of verifying the legitimacy of collaborators or users. This challenge is most important for organizations working with many different partners or users. To ensure equitable access, those organizations need systematic, objective, and transparent processes. Practices for verifying the legitimacy of users or customers in the life sciences have been explored in other contexts, including customer screening by synthetic nucleic acid providers,[51] and resources to support this type of decision-making are becoming more widely available.[52] Lessons can also be learned from screening procedures at other life sciences platforms, such as the U.K. Biobank,[53] the plasmid repository Addgene,[54] or the pathogen sequence database GISAID.[55] However, guidance and best practices specific to biological AI tools are lacking.

Data Governance

Biological AI tools depend on data, and all case studies highlighted the need for approaches for managing access to data. The link between data governance and governance of biological AI tools was made explicit by one case study interviewee, who stated that biological AI tools are simply a condensed form of the data that was used to train them. For some types of biological data, norms and policies that govern how they should be accessed exist already. For example, some databases that include pathogen genomes and variants (e.g., GISAID) require registration and an affiliation with a life sciences institution. Data related to humans are subject to a range of restrictions depending on how personally identifiable the data are. Export controls can also apply to certain databases that are shared internationally but are not publicly available.

Case study interviewees who worked with these types of data were familiar with many of these restrictions and were committed to respecting the rules. One interviewee described evaluating each dataset to determine whether it included sequences that might be found in pathogen databases or that might constitute existing intellectual property so that those sequences (or the whole dataset) could be excluded. This exclusion would enable the organization to share the data more openly. However, some organizations depend on a broad range of biological data to meet their missions, including some that are subject to access restrictions. Additional approaches are needed to properly federate or combine different types of data for training AI models and to manage access to tools and results responsibly.

Collection of Databases and Tools Generated by External Partners

Some organizations in the managed access case studies developed their own databases or biological AI tools, but nearly all organizations integrated databases or tools from external sources. In addition to following established policies and norms for biological data, those organizations that collect such resources from external partners highlighted the need for licensing agreements to ensure that the data or tools are used as intended by the original developer. Databases or developers may have differing requirements, and it can be challenging to ensure that each database or tool is accessed according to its specific licensing agreement. Multiple case study interviewees stated that standardized licensing agreements or templates would be helpful, both for establishing shared expectations across the community and for reducing the time and

resource burden on individual organizations to develop such agreements. Other important considerations include intellectual property ownership, authorship for publications, export controls, and potential legal liabilities related to how the resources are accessed or used.

Norms for Openness versus Managed Access

In the life sciences community, strong norms exist for openness to support transparency, peer review, reproducibility, and opportunities to build on previous work. Case study interviewees highlighted the tension between these norms and managed access for security purposes. One practical challenge for managed access is that most scientific journals and many funders require that data and tools be openly available to support publication of results, which limits publication pathways. Multiple case study interviewees described negotiations with publishers regarding which data or tools could be fully released and which should be more carefully managed. Ideas to support wider adoption of managed access included the following:

- Community engagement and biosecurity training materials to increase awareness of security concerns
- Development of standardized risk evaluation methods for biological resources to improve transparency, defensibility, and communication of decision-making related to managed access
- Support for platforms for managed access to biological AI tools and related resources that meet the needs of journals and other community gatekeepers

Appendix B. Managed Access Case Studies

1. Small, international organization with datasets and a biological AI tool it developed as part of a study on nucleic acid synthesis screening

The organization has a tool to generate sequences to test the vulnerability of widely used nucleic acid synthesis screening approaches to AI-generated sequences. The sequences are designed to yield protein structures similar to those known to pose pathogen or toxin risks but have been obfuscated in a way that reduces sequence homology to natural sequences. In addition to code for the tool itself, the organization has datasets of obfuscated sequences related to pathogens and toxins (including information on whether those sequences are likely to be functional) and of screening successes and failures. The organization would like to share these resources for publication purposes, to generate productive feedback on its study, and to further advance work related to biosecurity sequence screening. However, these resources pose biosecurity risks owing to information hazards associated with the datasets and the potential for misuse of the tool itself.

The organization is considering a tiered approach in which it would grant access to different tiers of resources based on whether it has verified the user's identity, it has verified that the user has some affiliation with the life sciences community, and the user has provided a reasonable explanation for why they want access to the materials. In addition, the user would have to sign a data use agreement that would depend on which of the resources they wanted to access (specific stipulations in these agreements are still under development).

Resources for managed access are a key challenge for the organization. In addition to staff time already spent, the organization anticipates that this approach will require some amount of time for each user to be verified. It also points to potential legal fees to ensure that data use agreements meet its needs and that export control considerations are addressed. To support managed access in this context, it would be helpful to have guidance or best practices for matching tiers of risk to appropriate levels of managed access, tools for efficient verification for user legitimacy, and templates or examples of data use or licensing agreements for sharing data.

2. Small foundation that enables sharing of AI-ready datasets for development of biological AI tools

The foundation collects existing biological datasets and provides financial support to develop new datasets to drive development of biological AI tools. Many existing datasets are open source, but newly developed ones are generally granted an embargo period to protect the ability of dataset developers to publish or file provisional patent applications using the data. The foundation works to ensure that these datasets do not pose significant biosecurity risks—for example, by manually curating them to remove sequences that match known pathogen or toxin sequences. The foundation also screens its datasets for sequences that could be considered protected intellectual property.

Managed access in this case study is primarily implemented to protect the publication embargo period for dataset contributors. Access is granted to partners who request access by email, set up a username and password, and use an assigned API key. To date, the foundation has worked primarily with individuals and groups it knows already, so it has not felt the need for more formal methods to verify user legitimacy.

In this case study, a key challenge for managed access is meeting the needs of the dataset developers, which often are not addressed by a one-size-fits-all licensing agreement. Developing an appropriate strategy took significant time and legal fees. The foundation also pointed to ongoing hosting fees, administrative challenges in ensuring that access is in compliance with the various requirements, and concerns about legal liabilities. It would be helpful to have more standardized templates, tools, and best practices for licensing agreements for biological resources. In the future, tools and best practices for verifying user legitimacy might be needed.

3. U.S. national laboratory with many types of biological resources developed for collaborations and applications in biodefense

In this case study, multiple types of resources support biodefense, including datasets, software tools, design systems with generative AI models, and integrated systems that incorporate experimental work to validate models. Many of these resources were developed in collaboration with academic and industry partners, and the laboratory anticipates that these tools will be useful for many different applications and users. To help determine biosecurity risks and appropriate levels of managed access, the laboratory organized an ad hoc consultation process with biosecurity experts, but it noted that there is no established methodology for these evaluations.

The suite of tools held by the laboratory is subject to different tiers of access. Some of the datasets and tools have been fully published. The design systems that include generative AI models will be shared with vetted collaborators such as academic labs, biotechnology companies, and others. For the integrated systems that include experimental data and feedback into generative AI models, access would require physical access to facilities, which is granted only with direct supervision of any external users. In the case study, funders generally make the determination about who should be vetted, and the laboratory itself is responsible for conducting due diligence to ensure the collaborators' legitimacy. The laboratory hopes and anticipates that it is well positioned to expand the number of partnerships to include many different collaborators and development and hosting of a wide range of biological resources, even beyond biodefense.

As a national laboratory, the organization did not experience any challenges related to resources or personnel for managed access, costs for computational infrastructure, verification of the legitimacy of users or collaborators, or establishment of licensing agreements. Instead, it pointed to broader challenges for managed access to biological AI tools and other resources, including strong cultural norms supporting openness. To support managed access, the laboratory believes it will be important to be more systematic and transparent about how risks and benefits are evaluated and how managed access judgments are made.

4. International collaboration with virology-specific databases and biological AI tools for public health applications

The organization provides funding and brings together partners and international collaborators to develop biological resources and systems for vaccine development. These capabilities include knowledge, databases, tools, and biological AI models relevant to high-priority pathogens and preparedness capabilities as well as integrated experimental systems for validating models and manufacturing of vaccines. The organization believes that many of these data and tools pose some biosecurity risk and it has worked with its collaborators to identify and assess risk, articulate principles to guide responsible conduct, and operationalize risk-mitigation approaches that can create global norms and best practices. However, there are no systematic ways to determine risks, and evaluations are complex with many factors to consider—for example, accidental and deliberate misuse of AI tools, data security, and incorporation of high-throughput experimental systems.

Managed access for the case study is multifaceted, and the organization is working to develop a comprehensive approach. It funds many international collaborators and plans to define expected behaviors and establish contractual requirements that could include responsible sharing of models, data, and other resources. For some future outputs, it anticipates many users from all over the world who will need to be vetted in some way. Given the organization's mission, it is important to ensure equitable and efficient access for all responsible users.

The biggest challenges for this case study are related to governance for a managed access approach, particularly in a multinational context. One challenge arises from the types of data incorporated into the organization's models. In addition to pathogen data and broader biological data, some of its models might incorporate human genomic data or data relevant to intellectual property, which may require a specialized approach. Some of its databases or models may also fall under export control rules, which can make it difficult to share. As a funder, the organization is also aware of legal liabilities that may arise in this complex landscape.

5. Funder and developer of biological AI tools with a platform to support discovery across the life sciences community

The case study focuses on an organization that supports the development of a variety of biological databases, tools, and foundation models relevant to health and disease. To better understand the risks associated with these resources, the organization has conducted detailed evaluations in consultation with biosecurity experts. Some of its smaller models have already been fully released, but it is considering a range of risk-mitigation options, including managed access, for any tools that pose significant risks.

A managed access framework for the organization has not yet been finalized, but it is considering the option to host some biological AI models and enable access through APIs. It is considering establishing criteria—such as requiring valid academic credentials—for access to these models, and it may set up different tiers so that some users have access through an API while others can access the source code, training data, or other model components. The organization is also working to establish a licensing agreement for all users that incorporates responsible use principles, and it hopes to develop training materials to increase awareness about biosecurity risks.

The organization pointed to significant challenges for managed access owing to a lack of guidance or best practices for risk assessment, matching levels of risk with appropriate mitigation approaches, and other implementation needs, in particular, for biological AI models. It is working to develop many resources itself, such as licensing agreements and biosecurity training materials, and it hopes to make them available to the broader life sciences community. In addition to these practical considerations, the organization also highlighted ongoing philosophical and cultural questions about how best to balance risk mitigation with the benefits of openness.

6. U.S. national laboratory with an AI tool developed for use only by U.S. federal government agencies

The group has developed a detailed, agent-based AI model to predict scenario-based epidemiological outcomes in the United States. Other epidemiological models have been openly published, but the group believes that its model poses particular biosecurity risks because it incorporates granular information specific to U.S. geography and population characteristics. In this case, the U.S. government sponsor provided straightforward guidance that access to this model should be restricted. In the absence of this guidance, the group believes it would have been much more difficult to determine how to appropriately manage access to the model and its outputs.

When a U.S. government agency requests access to information from this model, the group runs the model based on requested scenarios. The information is then provided to the agency in a series of briefings so that the agency has access to the required information, context, and expertise to interpret the model's outputs. Running the full model requires extensive computational resources, which causes a bottleneck that limits its use.

In this case study, the group did not experience significant challenges in resources, decision-making, or implementation of its managed access approach. However, the group highlighted the fact that its closed approach had serious implications for its ability to publish results and outcomes from the model. Although there are some journals that will allow publication of study outputs without requiring open access to the model, options are very limited. The group believes that a broader managed access framework or platform could be successful if it is systematic, standardized, and developed in collaboration with journals and other members of the scientific community.

Appendix C. Project Participants

More than 30 experts, including those listed in this appendix, contributed to this report. The list includes those interviewed for their experience and perspectives on managed access, participants in a September 2025 workshop previewing the report, and others who provided feedback and support through various channels. This report and its recommendations were developed by the authors and do not necessarily represent the views of all participants.

Tessa Alexanian
International Biosecurity and Biosafety Initiative for Science

Allison Berke, PhD
RAND

Polina Brangel, PhD
Coalition for Epidemic Preparedness Innovations

Jim Brase, PhD
Lawrence Livermore National Laboratory

Elizabeth (Beth) Cameron, PhD
Brown University Pandemic Center

Neil Cherian
Coalition for Epidemic Preparedness Innovations

Barbara Del Castello
RAND

Sara Del Valle, PhD
Los Alamos National Laboratory

Nicholas Generous
Los Alamos National Laboratory

Anthony Gitter, PhD
University of Wisconsin-Madison, Morgridge Institute for Research

Ben Gordon, PhD
Asimov

Björn Grüning, PhD
University of Freiburg

Dr. Moritz Hanke
Johns Hopkins University

Ian Haydon, PhD
University of Washington Institute for Protein Design

Andrew Hebbeler, PhD
Coalition for Epidemic Preparedness Innovations

Nathan Hillson, PhD
Lawrence Berkeley National Laboratory

Corey Hudson, PhD
The Align Foundation

Alex John London, PhD
Carnegie Mellon University

Sharon Malonza
Centre for Long-Term Resilience

Dr. Cassidy Nelson
Centre for Long-Term Resilience

Dr. Jassi Pannu
Johns Hopkins University

Girish Patangay
Chan Zuckerberg Initiative

Claire Qureshi
Sentinel Bio

Ryan Ritterson, PhD
Deloitte

Sebastian Rivera, PhD
Engineering Biology Research Consortium

Nicole Tensmeyer, PhD
Deloitte

Dr. Toby Webster
RAND

Nicole Wheeler, PhD
University of Birmingham

Jaime Yassif, PhD
Nuclear Threat Initiative

Sana Zakaria, PhD
RAND Europe
About the Authors

About the Authors

Sarah R. Carter, PhD is the Principal at Science Policy Consulting LLC. For more than 15 years, she has focused on advances in the tools and capabilities for engineering biology, biosecurity screening frameworks, and international norms for biosecurity. Her recent work, including projects with the Nuclear Threat Initiative (NTI), the Coalition for Epidemic Preparedness Innovations (CEPI), and other non-profit organizations, focuses on DNA synthesis screening, the implications of artificial intelligence (AI) for biosecurity, and approaches to reduce risks related to the misuse of biological AI tools. Previously, she worked in the Policy Center of the J. Craig Venter Institute and at the White House Office of Science and Technology Policy (OSTP). She is a former AAAS S&T Policy Fellow and a former Mirzayan S&T Policy Fellow of the National Academies. She earned her PhD from University of California, San Francisco, and her bachelor's degree from Duke University.

Greg Butchello is a program officer with NTI's Global Biological Policy and Programs team (NTI | bio). In this role, he supports many of the team's projects that seek to counter emerging biological threats, including the AIxBio Global Forum and multiple pilot projects to develop guardrails for biological AI tools. He also co-leads the Next Generation for Biosecurity portfolio and supports the team's congressional engagement. Prior to this, he served as the Executive Assistant and Conference Coordinator for both the NTI | bio team and the Nuclear Materials Security team. Mr. Butchello earned his MPA at American University, with a focus on domestic social policy. He also holds a BA in Art History and History from the University of Tulsa.

Endnotes

[1] Sarah R. Carter et al., *The Convergence of Artificial Intelligence and the Life Sciences* (Washington, DC: NTI, 2023), www.nti.org/analysis/articles/the-convergence-of-artificial-intelligence-and-the-life-sciences/; and Sarah R. Carter et al., *Developing Guardrails for AI Biodesign Tools* (Washington, DC: NTI, 2024), www.nti.org/analysis/articles/developing-guardrails-for-ai-biodesign-tools/.

[2] NASEM (National Academies of Sciences, Engineering, and Medicine), *The Age of AI in the Life Sciences: Benefits and Biosecurity Considerations* (Washington, DC: National Academies Press, 2025), doi.org/10.17226/28868.

[3] NTI (Nuclear Threat Initiative), "Statement on Biosecurity Risks at the Convergence of AI and the Life Sciences," July 17, 2025, www.nti.org/analysis/articles/statement-on-biosecurity-risks-at-the-convergence-of-ai-and-the-life-sciences/; and Yoshua Bengio et al., "International AI Safety Report," U.K. Government, Department of Science, Innovation, and Technology, January 2025, www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai.

[4] For example, see "Community Values, Guiding Principles, and Commitments for the Responsible Development of AI for Protein Design," Responsible AI x Biodesign, March 8, 2023, responsiblebiodesign.ai/.

[5] Carter et al., *Developing Guardrails for AI Biodesign Tools*; and NASEM, *Disseminating In Silico and Computational Biological Research: Navigating Benefits and Risks: Proceedings of a Workshop* (Washington, DC: National Academies Press, 2025), doi.org/10.17226/29174.

[6] Carter et al., *Developing Guardrails for AI Biodesign Tools*; and Mengdi Wang et al., "A Call for Built-in Biosecurity Safeguards for Generative AI Tools," *Nature Biotechnology* 43 (2025): 845–47, doi.org/10.1038/s41587-025-02650-8.

[7] For example, the U.K. Biobank has an application process that is overseen by an access committee to ensure that users are "bona fide researchers." U.K. Biobank, "Access Procedures, Version 2.1," July 2022, www.ukbiobank.ac.uk/wp-content/uploads/2025/01/Access-procedures.pdf.

[8] James Brian Byrd et al., "Responsible, Practical Genomic Data Sharing That Accelerates Research," *Nature Review Genetics* 10 (2021): 615–29, doi.org/10.1038/s41576-020-0257-5.

[9] GISAID (Global Initiative on Sharing All Influenza Data), "GISAID EpiFlu™ Database Access Agreement," March 16, 2011, gisaid.org/terms-of-use/.

[10] "AlphaFold Server," Google DeepMind, accessed October 22, 2025, alphafoldserver.com/.

[11] Ewen Callaway, "AI Protein-Prediction Tool AlphaFold3 Is Now More Open," *Nature*, November 11, 2024, www.nature.com/articles/d41586-024-03708-4.

[12] Thomas Hayes et al., "Simulating 500 Million Years of Evolution with a Language Model," *Science* 387, no. 6736 (2025): 850–58, doi.org/10.1126/science.ads0018.

[13] Conor Griffin et al., "Our Approach to Biosecurity for AlphaFold 3," Google DeepMind, accessed May 8, 2024, storage.googleapis.com/deepmind-media/DeepMind.com/Blog/alphafold-3-predicts-the-structure-and-interactions-of-all-lifes-molecules/Our-approach-to-biosecurity-for-AlphaFold-3-08052024; and Evolutionary Scale, "ESM3: Simulating 500 Million Years of Evolution with a Language Model," June 25, 2024, www.evolutionaryscale.ai/blog/esm3-release.

[14] NASEM, "*Disseminating In Silico and Computational Biological Research*."

[15] Eric Horvitz, "When AI Meets Biology: Promise, Risk, and Responsibility," Microsoft Research (blog), October 6, 2025, www.microsoft.com/en-us/research/blog/when-ai-meets-biology-promise-risk-and-responsibility/.

[16] Bruce J. Wittmann et al., "Strengthening Nucleic Acid Biosecurity Screening against Generative Protein Design Tools," *Science* 390, no. 6768 (2025): 82–87, doi.org/10.1126/science.adu8578.

[17] Zaixi Zhang et al., "FoldMark: Safeguarding Protein Structure Generative Models with Distributional and Evolutionary Watermarking" (preprint, submitted June 2025), doi.org/10.1101/2024.10.23.619960.

[18] Douglas Densmore et al., "A Proposal for Biodesign Metadata Exchange for Use in Biosecurity" (white paper, Nuclear Threat Initiative, Washington, DC, August 1, 2025), www.nti.org/analysis/articles/white-paper-a-proposal-for-biodesign-metadata-exchange-for-use-in-biosecurity/.

[19] Carter et al., *Developing Guardrails for AI Biodesign Tools*; and Wang et al., *A Call for Built-in Biosecurity Safeguards*.

[20] Irene Solaiman et al., "Beyond Release: Access Considerations for Generative AI Systems" (preprint, submitted February 2025), doi.org/10.48550/arXiv.2502.16701.

[21] "Galaxy," accessed October 22, 2025, usegalaxy.org/.

[22] "Exploring the MIT Open Source License: A Comprehensive Guide," MIT Technology Licensing Office, tlo.mit.edu/understand-ip/exploring-mit-open-source-license-comprehensive-guide.

[23] NTI, "Statement on Biosecurity Risks."

[24] Toby Webster et al., *Global Risk Index for AI-Enabled Biological Tools*, Appendix A.2 (Centre for Long-Term Resilience and RAND Europe, September 9, 2025), doi.org/10.71172/wjyw-6dyc.

[25] Richard Moulange et al., *Capability-Based Risk Assessment for AI-Enabled Biological Tools* (London, England: Centre for Long-Term Resilience, 2024), www.longtermresilience.org/reports/capability-based-risk-assessment-for-ai-enabled-biological-tools/; and Cassidy Nelson and Sophie Rose, *Understanding AI-Facilitated Biological Weapon Development* (London, England: Centre for Long-Term Resilience, 2023), www.longtermresilience.org/reports/understanding-risks-at-the-intersection-of-ai-and-bio/.

[26] Jaspreet Pannu et al., "Dual-Use Capabilities of Concern of Biological AI Models," *PLoS Computational Biology* 21, no. 5 (May 8, 2025), https://doi.org/10.1371/journal.pcbi.1012975.

[27] NASEM, *Age of AI in the Life Sciences*; and Allison Berke et al., *Data and AI-Enabled Biological Design: Risks Related to Biological Training Data and Opportunities for Governance* (RAND, June 30, 2025), www.rand.org/pubs/perspectives/PEA3886-1.html.

[28] Evolutionary Scale, "ESM Cambrian: Revealing the Mysteries of Proteins with Unsupervised Learning," December 24, 2024, www.evolutionaryscale.ai/blog/esm-cambrian#responsible-development; and CEPI (Coalition for Epidemic Preparedness Innovations), "Biosecurity Strategy Implementation Plan," accessed October 22, 2025, cepi.net/biosecurity.

[29] NASEM, *Biodefense in the Age of Synthetic Biology* (Washington, DC: National Academies Press, 2018), doi.org/10.17226/24890.

[30] National Research Council, *Biotechnology Research in an Age of Terrorism* (Washington, DC: National Academies Press, 2004), doi.org/10.17226/10827.

[31] NIH (National Institutes of Health), "US Government Releases Policy for Oversight of Dual Use Research of Concern and Pathogens with Enhanced Pandemic Potential," May 7, 2024, osp.od.nih.gov/us-government-releases-policy-for-oversight-of-dual-use-research-of-concern-and-pathogens-with-enhanced-pandemic-potential/.

[32] "Visibility Initiative for Responsible Science," Bio Policy & Leadership in Society, Stanford University, accessed October 22, 2025, biopolis.stanford.edu/virs.

[33] "Dual-Use Quickscan," Netherlands Biosecurity Office, dualusequickscan.com/en/.

[34] Webster et al., *Global Risk Index*, Appendix A.2.

[35] NASEM, *Age of AI in the Life Sciences*.

[36] "International Bio Funders Compact," Nuclear Threat Initiative, accessed October 22, 2025, www.nti.org/about/programs-projects/project/bio-funders-compact/; and NASEM, *Disseminating In Silico and Computational Biological Research*.

[37] NASEM, *Disseminating In Silico and Computational Biological Research*, figure 7.

[38] Tessa Alexanian and Sarah R. Carter, "Verifying Legitimacy: Findings from the Customer Screening Working Group, 2020–2023" (white paper, IBBIS, Geneva, February 2024), ibbis.bio/papers/whitepaper_2024_verifying_customer_legitimacy; Sarah R. Carter, *Developing a Customer Screening Framework for the Life Sciences* (Blueprint Biosecurity, March 2024), blueprintbiosecurity.org/works/kyc-report/; EBRC (Engineering Biology Research Consortium), *Strengthening a Safe and Secure Nucleic Acid Synthesis Ecosystem* (Emeryville, CA: EBRC, January 2025), doi.org/10.25498/E4311B; and Sarah R. Carter, Lucas Boldrini, and Tessa Alexanian, "Implementing Emerging Customer Screening Standards for Nucleic Acid Synthesis" (white paper, International Biosecurity and Biosafety Initiative for Science, Geneva, August 2025), ibbis.bio/wp-content/uploads/2025/11/IBBIS_Whitepaper_2025_Implementing-Emerging-Customer-Screening-Standards-for-Nucleic-Acid-Synthesis.pdf.

[39] IBBIS (International Biosecurity and Biosafety Initiative for Science), "Resources for Verifying the Legitimacy of Customers and Collaborators," accessed October 22, 2025, ibbis.bio/our-work/customer-screening/.

[40] CEPI (Coalition for Epidemic Preparedness Innovations), "Biosecurity Strategy Implementation Plan," April 2025, cepi.net/biosecurity.

[41] Examples of platforms that provide AI-ready databases include Therapeutics Data Commons (tdcommons.ai) and the Align Foundation (alignbio.org/).

[42] "Gated Models," Hugging Face, accessed October 22, 2025, huggingface.co/docs/hub/models-gated.

[43] Neurosnap, accessed October 22, 2025, neurosnap.ai/.

[44] Rowan, accessed October 22, 2025, rowansci.com/.

[45] "All Tools," Tamarind Bio, accessed October 22, 2025, app.tamarind.bio/.

[46] "Solutions," Levitate Bio, accessed October 22, 2025, levitate.bio/products.

[47] "Models," NVIDIA, accessed October 22, 2025, build.nvidia.com/models.

[48] Teselagen, accessed October 22, 2025, teselagen.com/.

[49] Benchling, accessed October 22, 2025, www.benchling.com/.

[50] Carter et al., *Developing Guardrails*.

[51] Alexanian and Carter, *Verifying Legitimacy*; Carter, *Developing a Customer Screening Framework*; and EBRC, *Strengthening a Safe and Secure Nucleic Acid Synthesis Ecosystem*.

[52] IBBIS, "Resources for Verifying the Legitimacy of Customers and Collaborators."

[53] U.K. Biobank, "Access Procedures, Version 2.1."

[54] Addgene Help Center, "Why Is Addgene 'Verifying' My Account and How Long Will It Take?," accessed October 22, 2025, help.addgene.org/hc/en-us/articles/206130535-Why-is-Addgene-verifying-my-account-and-how-long-will-it-take.

[55] GISAID, "GISAID EpiFlu™ Database Access Agreement."

About NTI

The Nuclear Threat Initiative (NTI) is a nonprofit, nonpartisan global security organization focused on reducing nuclear, biological, and emerging technology threats imperiling humanity. The biosecurity mission is conducted by NTI's Global Biological Policy and Programs (NTI | bio).

1776 Eye Street, NW | Suite 1000 | Washington, DC 20006 | **www.nti.org**

facebook.com/nti.org @NTI_WMD NTI_WMD Nuclear Threat Initiative