**Testimony of** 

Tom Inglesby, MD

Director, Johns Hopkins Center for Health Security Bloomberg School of Public Health

Before the

US Senate Committee on Health, Education, Labor and Pensions Subcommittee on Primary Health & Retirement Security

"Avoiding a Cautionary Tale: Policy Considerations for Artificial Intelligence in Health Care"

November 8, 2023

#### **Introduction**

Chairman Markey, Ranking Member Marshall, and distinguished members of the Committee, it is my pleasure to appear before you today to discuss the potential benefits and challenges related to artificial intelligence (AI) use in health care and public health. In order to harness the great promise that AI holds for benefits in health care and public health, AI risks (including privacy, data integrity, and bias) all need to be rigorously addressed.

Within the realm of AI models working in the biological sciences, I want to urge this Committee to place high priority on establishing strong governance over the highest potential dual-use risks of AI and biosecurity (AIxBio), which I judge to be: (1) the potential for AI to accelerate or simplify the reintroduction of particularly dangerous extinct viruses or dangerous viruses that only exist now within research labs; and (2) the potential for AI to enable, accelerate, or simplify the creation of entirely new biological constructs that could start a new pandemic. Taken together, AI foundation models like large language models (LLMs), and AI biological design tools (BDTs), such as models focused on protein design or immune evasion, could now or in the foreseeable future be misused to purposefully create such threats. We should start working to guard against these risks today.

My name is Tom Inglesby. I am Director of the Johns Hopkins Center for Health Security and Professor in the Department of Environmental Health and Engineering in the Johns Hopkins Bloomberg School of Public Health, with a Joint Appointment in the Johns Hopkins School of Medicine. I'm also a medical doctor with a background caring for patients with HIV, and I worked on the COVID pandemic response, including on resolving challenges around access to diagnostic testing for COVID. The opinions expressed herein are my own and do not necessarily reflect the views of Johns Hopkins University.

For 25 years, our Center's mission has been to protect people's health from major epidemics and disasters and build resilience to those challenges. Our Center is comprised of researchers and experts in science, medicine, public health, law, social sciences, economics, and national security – all focused on our mission to protect people's health from epidemics and disasters and ensure that communities are resilient to major challenges. Our team conducts independent research and analyzes how scientific and technological innovations can strengthen health security. Our Center founded the bipartisan Capitol Hill Steering Committee on Pandemic Preparedness and Health Security in 2020, in collaboration with Members of the House and Senate, as well as former Administration officials, as an educational forum to discuss new topics, technologies, and ideas that can improve domestic health security now and in the future. The Steering Committee has held over 20 sessions in the last three years intended to be of value to congressional offices working on pandemic and biosecurity challenges.

Today, I was asked to provide comments on how we can guard against potential harms of AI while at the same time working to ensure that AI, where implemented, is done so in ways that will improve patient experience and outcomes. In my testimony below, I provide my views on the enormous potential benefits of AI in health care and the substantial potential risks that need to be addressed before and while realizing those benefits. Prior to offering those views, I want to give my top line recommendations as to what Congress should be doing at this time to address the greatest AIxBio risks.

To that end, I recommend that Congress now build on the strong foundation provided by the October 30 Executive Order titled: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (EO #14110). I recommend that congressional actions related to this include:

- (1) Providing the Department of Health and Human Services (HHS) with the authority and resources to require anyone purchasing synthesized nucleic acids, regardless of the funding source, to purchase only from a provider or manufacturer that screens both orders and customers in a way that reduces the highest potential dual-use risks of AIxBio.<sup>1</sup>
- (2) Commissioning a rapid risk assessment to identify whether EO #14110 as written will adequately address high-end biological risks or whether congressional action is needed in the near-term to ensure prevention of those threats.
- (3) Requiring entities developing models with significant dual-use risks to red-team and evaluate their models, and task an agency with: (1) auditing those models; and (2) submitting a report to Congress with recommendations for new authorities that will be needed by the agency to take any appropriate remedial action should red-teaming, evaluations, or audits fail.

If taken now, these measures will reduce the risk of malicious and consequential misuse of AI-enabled biology while allowing AI developers and scientists to pursue beneficial uses of AI to improve the human condition.

## Medical and Public Health Benefits of AI and Recognition of Other Risks in Health Care

AI holds great promise for benefits in health care and public health. Potential benefits include earlier disease diagnoses, allowing doctors to intervene earlier in the course of an illness; reduced medical errors; more efficient or less invasive surgeries; lowering of administrative burdens on clinicians to allow more time with patients; and faster response times to patient questions. Researchers and companies may be able to create or use AI tools to help them accelerate development of vaccines and medicines and to significantly advance personalized medicine. AI may be able to improve disease surveillance and perhaps even provide earlier indicators of new outbreaks or epidemics. It will place stronger diagnostic and clinical tools in the hands of providers in the field or those in clinics far from more advanced health care systems.<sup>2</sup> AI could also assist with more careful monitoring of drug safety and help to improve, and potentially greatly accelerate, clinical trials of new medicines.

To realize these benefits, policymakers, companies, and health systems will need to take great care in implementing consequential AI systems, and all parties will need to address a series of risks and potentially serious challenges. For instance, developers could inadvertently introduce biases into the models that are being developed in AI health care systems. Policymakers and firms will need to ensure that privacy is protected so that individual patient information is not inappropriately accessed or shared publicly. This includes addressing cybersecurity issues in AI, such as the potential for offensive cyberAI to outstrip cyberAI's defensive capabilities, using lessons learned from cyber governance.<sup>3</sup> The quality and integrity of

<sup>&</sup>lt;sup>1</sup> (requiring that all federally funded entities conducting life-sciences research purchase synthetic nucleic acids only from providers or manufacturers that adhere to the screening framework developed by NIST). *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, 88 Fed. Reg. 75191 (Nov. 1, 2023), § 4.4(b)(iii).

<sup>&</sup>lt;sup>2</sup> World Health Organization (WHO), *Ethics and Governance of Artificial Intelligence for Health*, WHO (June 28, 2021), <u>https://www.who.int/publications/i/item/9789240029200</u>; IBM Education, *How Can Artificial Intelligence Benefit Healthcare*?, IBM (July 11, 2023), <u>https://www.ibm.com/blog/the-benefits-of-ai-in-healthcare</u>/.

<sup>&</sup>lt;sup>3</sup> Louis Columbus, *Defensive Vs. Offensive AI: Why Security Teams are Losing the AI War*, VENTUREBEAT (Jan. 3, 2023, 10:07 AM), <u>https://venturebeat.com/security/defensive-vs-offensive-ai-why-security-teams-are-losing-the-ai-war/</u>.

the training data for AI systems will need to be high – inaccuracies or skews in the data that AI systems are being trained on could lead to inaccurate or misleading results that could be damaging and hard to detect.<sup>4</sup>

There are additional legal and ethical risks associated with AI. When implementing the technology, it will be vital to ensure that AI is not used as a substitute for investment in and development of core health functions.<sup>5</sup> Many have identified these and other challenges, and it's good to see that US-based companies are trying to work with the government to find feasible ways of effectively mitigating the range of potential AI risks to health care. It will be important for Congress to regularly assess the extent to which AI developers and health care systems are addressing these risks, and to consider legislative remedies to address any clear gaps.

## The Need for Strong AIxBio Governance

One area of risk that deserves special and immediate attention is the potential for AI systems to create highconsequence biosecurity and biosafety risks. Leaders from the AI technology field have identified those risks as among their highest priority concerns, as have government officials and outside research groups focused on the establishment of AI governance systems.<sup>6</sup>

Signed last week, EO #14110 represents the strongest action on AI that any government has taken thus far. It sets out a series of high-level principles and priorities that broadly commit the country's AI path to: developing safe and secure AI systems; responsible innovation and competition; a commitment to supporting workers; advancing equity around AI; the protection of privacy and civil liberties; responsible federal use of AI; and strong global leadership.

As part of this overall approach, the EO identifies a series of specific risks the executive branch will work to address, including the risk that AI systems could substantially lower the barrier of entry to design, synthesize, acquire, or use biological weapons. It details a series of important steps the executive branch will take in the months ahead to develop guidance, identify new industry norms, and evaluate potential risks in order to protect against AI being deliberately misused for this purpose.

The EO directs the National Institute of Standards and Technology (NIST) to develop guidelines and best practices, with the aim of promoting consensus industry standards for safe and secure systems that include benchmarks for evaluating and auditing AI capabilities to cause harm, as well as guidance for AI developers regarding red-teaming practices and testing processes and environments. It also directs the Department of

<sup>6</sup> See, e.g., Diane Bartz, US Senators Express Bipartisan Alarm About AI, Focusing on Biological Attack, REUTERS (July 25, 2023,10:23 PM), <u>https://www.reuters.com/technology/us-senators-express-bipartisan-alarm-about-ai-focusing-biological-attack-2023-07-25/;</u> Congresswoman Anna G. Eshoo, *Eshoo Urges NSA & OSTP to Address Biosecurity Risks Caused by AI,* CONGRESSWOMAN ANNA G. ESHOO (Oct. 25, 2022), <u>https://eshoo.house.gov/media/press-releases/eshoo-urges-nsa-ostp-address-biosecurity-risks-caused-ai;</u> The White House, *Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence*, WHITE HOUSE (Oct. 30, 2023), <u>https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/; Nuclear Threat Initiative (NTI), *Report Launch: The Convergence of Artificial Intelligence and the Life Sciences*, NTI (Oct. 30, 2023), <u>https://www.nti.org/events/report-launch-the-convergence-of-artificial-intelligence-and-the-life-sciences/.</u></u>

<sup>&</sup>lt;sup>4</sup> World Health Organization (WHO), *Ethics and Governance of Artificial Intelligence for Health*, WHO (June 28, 2021), <u>https://www.who.int/publications/i/item/9789240029200</u>.

<sup>&</sup>lt;sup>5</sup> World Health Organization (WHO), *WHO Issues First Global Report on Artificial Intelligence (AI) in Health and Guiding Principles for Its Design and Use*, WHO (June 28, 2021), <u>https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use.</u>

Energy to implement tools and testbeds for evaluating AIxBio capabilities and to develop guardrails that reduce these risks.

The EO directs the Department of Commerce to require companies with frontier dual-use foundation AI models (models that could potentially lower barriers for designing/synthesizing bioweapons) to report activities related to the production of those models, the protection of key model characteristics, and the results of red-teaming tests.

The EO also directs the Office of Science and Technology Policy (OSTP) to establish a framework that encourages providers of synthetic nucleic acid sequences to implement comprehensive nucleic acid procurement screening mechanisms. As part of that effort, OSTP will need to establish criteria and mechanisms for identifying sequences that pose a risk to national security and determine methodologies for verifying performance of screening, including customer screening approaches. Six months after the creation of this framework, all agencies that fund life sciences work will establish that their funding recipients procure nucleic acid sequences from manufacturers that adhere to this framework.

My Center, along with other biosecurity-focused researchers and experts, as well as industry leaders from the companies that conduct nucleic acid synthesis, have been calling for the development of a framework to require those who procure nucleic acid sequences to purchase them from companies that are verified to be carefully screening orders and customers in order to deter and detect any potentially malicious actors. I'm very glad that the EO makes progress on this issue for those entities receiving federal funding.

I believe that this series of EO actions, taken together, are appropriate, important, strong actions that are needed to better assess, evaluate, test for, and diminish biological risks posed by new AI models. AI foundation models, LLMs, and AI biological design tools – such as those that help to design and predict structures of proteins, design viral vectors, or predict the properties of pathogens, host-pathogen interactions, or immune-system evasion – could be misused by accelerating the synthesis/manufacture of extinct or eradicated highly transmissible viruses, or by helping to design novel biological constructs capable of epidemic or pandemic spread. While more evaluation and study of these risks are clearly needed, preliminary evidence suggests that AI models could in the foreseeable future accelerate, simplify, or enable the creation of these risks. Early technical studies from nongovernmental research teams that I've been briefed on are quite worrying. As these assessments are ongoing, we need a governance process that will address risks identified during red-teaming exercises and other evaluations.

Beyond this EO, I have been encouraged by other developments to address these risks. I highly commend many of the AI companies for making voluntary commitments to pre-release internal and external security testing of their AI systems, which includes testing by independent experts to guard against biosecurity risks.<sup>7</sup> The first step in addressing risk is to identify it, and many of the companies developing frontier models have made progress in the past year in trying to understand the biosecurity risks that their models may pose and addressing those risks.<sup>8</sup>

https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/.

<sup>8</sup> See, e.g., Diane Bartz, US Senators Express Bipartisan Alarm About AI, Focusing on Biological Attack, REUTERS (July 25, 2023,10:23 PM), <u>https://www.reuters.com/technology/us-senators-express-bipartisan-alarm-about-ai-focusing-biological-</u>

<sup>&</sup>lt;sup>7</sup> The White House, *Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI*, WHITE HOUSE (July 21, 2023), https://www.whitehouse.com/statements.releases/2023/07/21/fact sheet hiden harris administration secure

I'm also encouraged by the Institute for Protein Design's community-wide effort to develop new voluntary guidelines for researchers to follow as they apply AI to protein research. Such commitments can help establish community standards and encourage ethical behavior on the part of individual scientists by, for example, creating an obligation to report any concerning research practices.<sup>9</sup>

Strong governance will also require international collaboration. That is why I'm very pleased to see that the US and 27 other countries recognized the special risks that AI poses in biotechnology in the recently signed Bletchley Declaration by Countries Attending the AI Safety Summit.<sup>10</sup> I'm further encouraged that at least two Artificial Intelligence Safety Institutes have already been stood up – one in the UK and one at NIST in the US Department of Commerce – to provide testing environments for researchers to evaluate emerging AI risks, such as those at the intersection of AI and biotechnology.

#### **Recommendations**

Congress should ensure that as the US government acts to mitigate the risks of AIxBio, it set as its highest priority the reduction of the two most consequential biological risks, which I argue are: (1) the potential for AI to accelerate or simplify the reintroduction of particularly dangerous extinct viruses or dangerous viruses that only exist now within research labs; and (2) the potential for AI to enable, accelerate, or simplify the creation of entirely new biological constructs that could start a pandemic.

While I am encouraged by recent actions being taken by the US government, industry developers of powerful AI technologies, and researchers in the field, there are series of steps that I think will be important for Congress to attend to in the time ahead to ensure that these two most consequential biological risks are addressed. They include:

# (1) Providing HHS with the authority and resources to require anyone purchasing synthesized nucleic acids, regardless of the funding source, to purchase only from a provider or manufacturer that screens both orders and customers in a way that reduces the highest potential dual-use risks of AIxBio.

Our increasing ability to automate scientific experiments, cheaply synthesize nucleic acids, and autonomously generate biological constructs will likely speed up development of drugs and devices to protect and prolong human health and allow the advent of enormously powerful medical tools that will protect millions of American lives, such as personalized medicine.<sup>11</sup> But we must ensure at the same time that these new powers are not used maliciously to cause great harm. Certain AI models will likely help to accelerate the transition across the "digital-to-physical" boundary – they may also enable digitally designed threats to turn into physical biological risk. They could be used to help malicious actors create highly dangerous and transmissible pathogens. Without a strong screening framework in place and required of all

attack-2023-07-25/ (Anthropic warning Senators about biological risks during congressional testimony); Anthropic, *Frontier Threats Red Teaming for AI Safety*, ANTHROPIC (July 26, 2023), <u>https://www.anthropic.com/index/frontier-threats-red-teaming-for-ai-safety</u> (Anthropic developing red-teaming tests to guard against biosecurity risks).

<sup>&</sup>lt;sup>9</sup> Institute for Protein Design (IPD), *Results from our Summit on Responsible AI*, IPD (Oct. 31, 2023), <u>https://www.ipd.uw.edu/2023/10/responsible-ai-summit/</u>.

<sup>&</sup>lt;sup>10</sup> The Prime Minister's Office, *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023*, PRIME MINISTER'S OFFICE (Nov. 1, 2023), <u>https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.</u>

<sup>&</sup>lt;sup>11</sup> Kanika Jain, *Synthetic Biology and Personalized Medicine*, 22 MED. PRINC. PRAC. 209 (2013), https://doi.org/10.1159/000341794.

companies, such actors could exploit companies that do not screen customers or orders, or they could find gaps in screening programs that are weak or insufficient to guard against exploitation.<sup>12</sup>

In order to secure the digital-to-physical frontier, it will be critical to implement mandatory screening policies for gene synthesis providers and manufacturers. EO #14110 requires that all federally funded entities conducting life sciences research must purchase synthetic nucleic acids from gene synthesis providers or manufacturers that adhere to a gene synthesis screening framework to be developed by OSTP.<sup>13</sup> This is an excellent initial step, but Congress should further provide HHS – as by far the largest government funder of life sciences research – with the authority and resources to expand this requirement to <u>all</u> US purchasers of synthetic nucleic acids, not just those receiving federal funding. There is broad public support for this – a recent poll found that 61% of Americans of all political affiliations support such an expansion, while only 12% do not.<sup>14</sup> My understanding is that the EO's screening requirements were applied only to federally funded entities because the authority to regulate the purchases by other entities in this manner does not currently exist within the executive branch. That suggests that action by Congress is vital. Congress should also give HHS the authority and resources to set up verification mechanisms to ensure that manufacturers and purchasers comply with screening requirements.

While Congress works to ensure that US gene synthesis providers follow OSTP's framework, the executive branch should focus on promoting the adoption of similar standards internationally. Around 60% of the gene synthesis market sits outside of North America.<sup>15</sup> Not only does this mean that malicious actors within the US can access international providers, but as COVID-19 demonstrated, borders are not a protection against disease – a gene synthesis-driven outbreak abroad could have terrible impact in the US. It is therefore crucial that the executive branch works to create a widely adopted international agreement that requires all gene synthesis providers globally to adhere to rigorous screening standards. The framework that will be developed as part of this EO will provide a vital starting point for such an agreement.

# (2) Commissioning a rapid report to identify whether EO #14110 as written will adequately address high-end biological risks or whether congressional action is needed in the near term to prevent those threats.

Although EO #14110 requires studies and reports on AIxBio risks,<sup>16</sup> those studies and reports (1) are not required to be reported to Congress; (2) will not include any new legislative recommendations; and (3) do not clearly prioritize high-end biological risks.

For example, the EO requires the Department of Homeland Security (DHS) to submit a report to the president on the potential for AI to be misused to enable the development or production of chemical, biological, radiological, and nuclear (CBRN) threats. It also requires the Department of Defense (DOD) to commission a report on biosecurity risks from AI. These are important actions for the executive branch to take. However, given the fast-moving nature of this technology and Congress's role in ensuring that the

<sup>&</sup>lt;sup>12</sup> The Hon. Mark Dybul et al., *Biosecurity in the Age of AI: Chairperson's Statement*, HELENA (July 2023), <u>https://www.helenabiosecurity.org.</u>

<sup>&</sup>lt;sup>13</sup> § 4.4(b)(iii).

<sup>&</sup>lt;sup>14</sup> Artificial Intelligence Policy Institute (AIPI), *Vast Majority of US voters of All Political Affiliations Support President Biden's Executive Order on AI*, AIPI (Oct. 30, 2023), <u>https://theaipi.org/poll-biden-ai-executive-order-10-30/</u>.

<sup>&</sup>lt;sup>15</sup> (though the market share of the US is expected to increase in coming years). Global Market Insights (GMI), *Gene* Synthesis Market - By Method (Solid-phase Synthesis), By Services (Antibody DNA Synthesis), By Application (Vaccine Development) By End-use (Academic and Research Institutes, Biopharmaceutical Companies,) & Forecast 2023 – 2032, GMI (May 2023), <u>https://www.gminsights.com/industry-analysis/gene-synthesis-market.</u>

<sup>&</sup>lt;sup>16</sup> §§ 4.4(a), 4.6.

executive branch has the tools and resources it needs to appropriately govern, Congress should commission a rapid report to identify whether EO #14110 as written will adequately address high-end biological risks or whether congressional action is needed in the near term to ensure prevention of those threats.

The need for this focus on high-end risks is akin to the important focus that is warranted around the governance of enhanced potential pandemic pathogen (ePPP) research. The US government should carefully scrutinize research that can reasonably be anticipated to create novel pandemic threats, lest we face the devastating consequences of an accident or deliberate misuse. Similarly, we should advance cautiously – and with full awareness of the relevant risks – as we fund and promote the creation of advanced AI models. In prior work on other issues related to biological threats, I have seen efforts that have neglected or paid insufficient attention to high-end biological risks, and I fear that the same thing could happen in this context.

Commissiong a rapid report on high-end biological risks posed by AI would provide timely clarity to Congress as it considers how to ensure the country is harnessing the incredible transformative power that AI promises in health care, public health, and broader society while guarding against its greatest risks. It would be logical for the Administration for Strategic Preparedness and Response (ASPR) to have responsibility for such a report given its responsibilities around genome synthesis screening and assessment of risks related to ePPP research.

# (3) Requiring entities developing models with significant dual-use risks to red-team and evaluate their models, and task an agency with: (1) auditing those models; and (2) submitting a report to Congress with recommendations for new authorities that will be needed by the agency to take any appropriate remedial action should red-teaming, evaluations, or audits fail.

Just as EO #14110 establishes a safety program at HHS that provides for remedial action if it finds harms or unsafe health care practices involving AI,<sup>17</sup> so too should Congress establish a program that provides for remedial action in the event that red-teamers demonstrate AI models enable high-end biological risks, evaluations identify high-end biological risks, or audits find that a company did not provide accurate information regarding high-end biological risks. What is currently required by the EO in the area of high-end biological risks is that companies developing or intending to develop dual-use foundation models must report relevant technical information to the federal government, including red-teaming performance related to AIxBio risks.<sup>18</sup> However, the question that Congress should address is: what happens in the event of failures? What can the government do if tests show that a model is too dangerous to release safely?

EO #14110 does not actually require companies to conduct red-teaming tests, evaluations, or audits. Instead, the EO simply requires that if a company voluntarily opts to red-team its dual-use foundation model, the results of those tests must be reported. <sup>19</sup> Moreover, the EO does not require individuals or groups that may develop AI systems in the future to report the same activities required of companies in the EO.<sup>20</sup>

<sup>&</sup>lt;sup>17</sup> The White House, *Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence*, WHITE HOUSE (Oct. 30, 2023), <u>https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.</u> <sup>18</sup> § 4.2(i).

 $<sup>^{19}</sup>$  *Id*.

<sup>&</sup>lt;sup>20</sup> Compare § 4.2(i) with § 4.2(ii). I suspect that this is because individuals or groups, such as academic institutions, are not currently developing frontier AI models. However, this could shift in the future, such as if the National AI Research Resource (NAIRR) provides independent AI researchers and students with significantly expanded access to computational resources. Accordingly, a capabilities-based requirement rather than an entity-based requirement seems warranted.

Accordingly, Congress should develop legislation to require all entities (not just companies) developing models with high-end, dual-use biological risks<sup>21</sup> to red-team, evaluate, and audit their models.

Additionally, while NIST is tasked with developing auditing standards in the EO, it's unclear whether any US government agency would have the authority to require entities to grant the government permission to audit those models, by which I mean the assessment of developers' red-teaming efforts as well as an evaluation of frontier models by the government itself. Nor is it clear by what authority the US government could take remedial action should its evaluation, or that of the developers, find a model dangerous. Congress should therefore task an agency with: (1) auditing those models as described above, as the agency deems necessary; and (2) submitting a report to Congress with recommendations for new authorities that will be needed by the agency to take any appropriate remedial action such as pausing development until safety measures can be implemented, cessation of development, or directing the developer to face other consequences if red-teaming, evaluations, or audits fail. In conducting these evaluations, agencies should of course consider both the most extreme risks posed by advanced models as well as their potential benefits, both in detecting and flagging pandemic threats and in mitigating them through vaccine and drug design.

One of the most concerning risks of AI models is that if they become wholly open source and available on the internet, they cannot be recalled.<sup>22</sup> That is why red-teaming, evaluations, and audits will be so important to conduct before future dual-use, high-end risk bio models are made open source – we will only have one chance to get it right for each release.

It will also be important for Congress to consider how to support the development of a skilled workforce able to sufficiently red-team frontier dual-use foundation models for the highest-consequence biological risks. Providing these authorities will ensure that the AI systems that could be used to design new effective pharmaceuticals, make breakthroughs in fundamental biology, and give doctors powerful new diagnostic tools do not create new pandemic risks that both endanger the public and threaten to undermine AI's great potential benefit.

### **Conclusion**

In order to harness the great promise that AI holds for benefits in health care and public health, AI risks (including privacy, data integrity, bias) will all need to be rigorously addressed. Within the realm of AI models working in the biological sciences, there are two high-consequence risks that deserve top priority for attention and strong governance: (1) the potential for AI to accelerate or simplify the reintroduction of particularly dangerous extinct viruses or dangerous viruses that only exist now within research labs; and (2) the potential for AI to enable, accelerate, or simplify the creation of entirely new biological constructs that could start a new pandemic.

While I am encouraged by recent actions taken by the US government, industry developers of powerful AI technologies, and researchers in the field, I outline above three steps that I think will be important for Congress to attend to in the time ahead to ensure that these high-consequence risks are addressed. If taken now, these measures will help to reduce the risk of malicious and consequential misuse of AI-enabled biology while allowing AI developers and scientists to pursue beneficial uses of AI to broadly improve medicine, public health, and patient outcomes.

<sup>&</sup>lt;sup>21</sup> Potentially subject to be defined by the actions taken in the EO. See § 4.2(b).

<sup>&</sup>lt;sup>22</sup> See, e.g., the leak of Meta's Llama model.