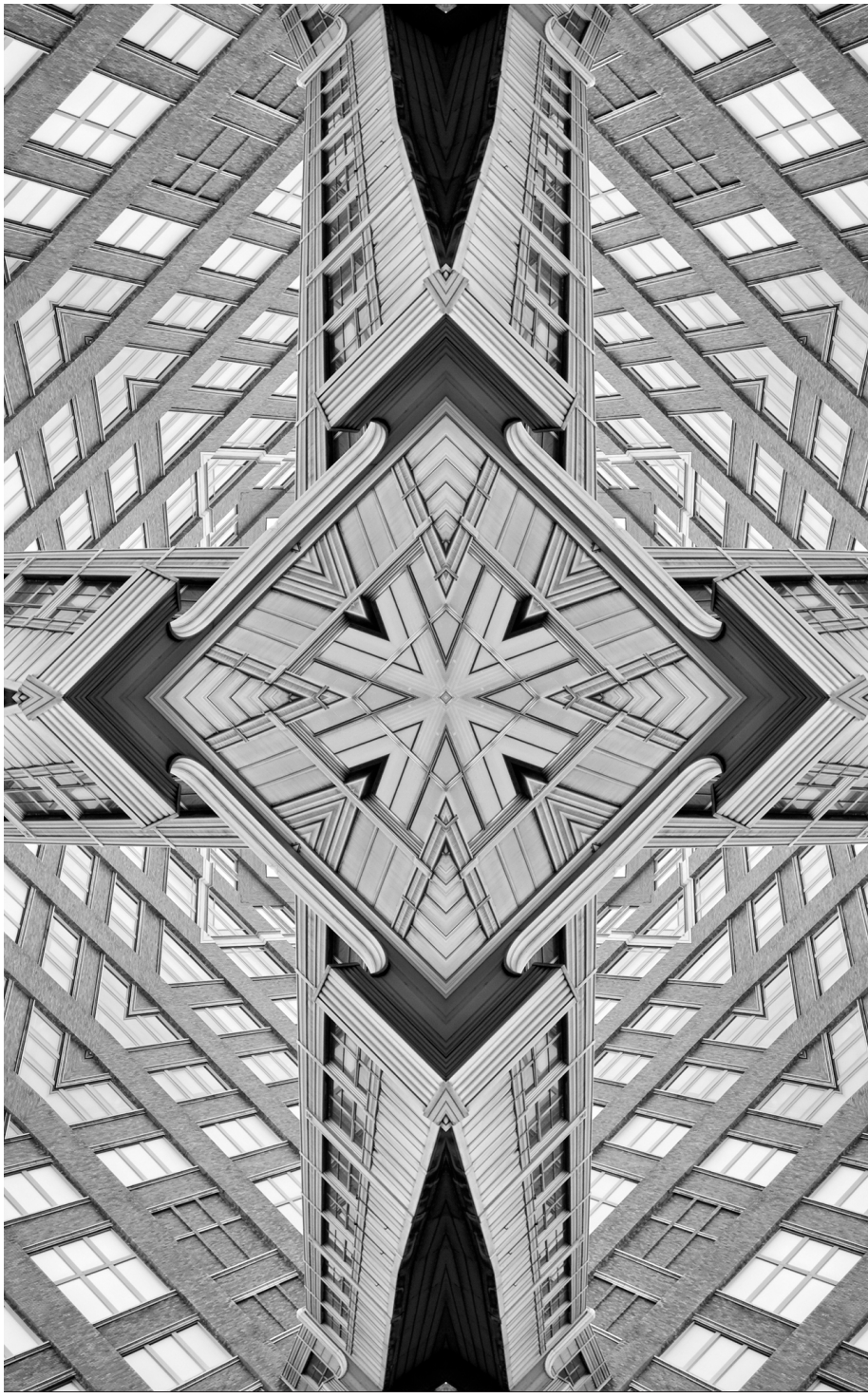
OBSERVER RESEARCH FOUNDATION

Issue Brief

ISSUE NO. 766
DECEMBER 2024



A Framework for Effective Risk Assessment of AI-Biotechnology Convergence

Lakshmy Ramakrishnan

Advances in disruptive technologies have enabled scientists to engineer biological systems and create more efficient and sustainable products for a range of purposes. The convergence of artificial intelligence (AI) with biotechnology has contributed to the development of therapeutics and vaccines, helped address food security, aided in generating innovative processes to mitigate the effects of climate change, and expanded the bioeconomy. However, there are concerns that AI-biotechnology tools can be misused to create harm, with catastrophic global consequences. This brief advocates for a nuanced approach to assess the risks associated with the convergence of AI with biotechnology and its potential for misuse and recommends a coordinated strategy.

Attribution: Lakshmy Ramakrishnan, "A Framework for Effective Risk Assessment of AI-Biotechnology Convergence," *Issue Brief No. 766*, December 2024, Observer Research Foundation.

Introduction

The convergence of artificial intelligence (AI) with biotechnology has ushered in an era of innovation in drug development, genomics, and diagnostics.[1] AI provides immense scope for interdisciplinarity owing to its ability to "integrate and analyse diverse data sets" from distinct domains.[2] For example, AI has aided in the development of new biomarkers for diagnosing and monitoring the treatment progression of Alzheimer's disease, the advancement of precision medicine for genetic disorders, and the production of novel biomolecules.[3] Furthermore, AI-biotechnology (AI-bio) tools have the potential to contribute to health security, particularly through their ability to predict public health threats.[4]

However, according to the UN's *Governing AI for Humanity* report, the use of AI needs to be governed because "no one currently understands all of AI's inner workings enough to fully control its outputs or predict its evolution."[5] Additionally, the report recommends that mechanisms need to be developed to prevent the misuse of AI.[6] Recent media reports[a,7] on the use of AI to help create pathogens and subsequent policy discussions led to the United Kingdom (UK) and the United States (US) acknowledging the potential misuse of AI by malicious actors to create bioweapons.[8] In this backdrop, there is a need to evaluate the real risk of AI-bio capabilities in the Indian context. While AI, in its current form, is unlikely to catalyse the development of biological weapons,[9] policymakers need to be aware of the scope, limitations, and feasibility of this potentiality to ensure effective safeguards against the misuse of these technologies in the future.

a A piece in the *New York Times* argued that "AI may save us or construct viruses to kill us," while *Science* and *Foreign Policy* posed the question whether chatbots could design bioweapons.

Historically, the use of bioweapons and bioterrorism has been rare. However, the dual-use capabilities of both biotechnology and AI have raised concerns about their malign capabilities.[10] Additionally, it is essential to balance the publicity around these technologies with the real risk they pose.

In their analysis of al-Qaeda's bioweapons strategy, US intelligence found that the US's policy of dissuasion, which involved insisting that bioweapons could be made cheaply and easily, drew the attention of al-Qaeda[11] and was a crucial factor that instigated al-Qaeda to attempt to make a bioweapon between 1997 and 2001.

In another instance, OpenAI conducted a stress test in 2023 to determine GPT-4o's ability to create biological threats.[12] The system card found that the model is able to "generate publicly accessible but difficult-to-find information, shortening the time users spend on research and compiling this information in a way that is understandable to a non-expert user."[13] This indicates that ChatGPT may be useful for those who "do not have access to formal scientific training", and this democratisation can increase the know-how of malevolent actors.[14] For example, it provided information on how to order synthetic oligonucleotides[b] from DNA companies, detailed laboratory protocols, and tips on how to troubleshoot experiments.

Around the same time, students at the Massachusetts Institute of Technology (MIT) conducted experiments with large language models (LLMs), including ChatGPT, to determine whether a virus with pandemic potential could be designed and ordered.[15] The chatbot provided information on four pathogens with pandemic potential, suggested mutations to enhance transmission, provided details on DNA technology companies that were unlikely to screen orders for potential misuse, and recommended contract research organisations to implement the protocols if the users lacked life-science skills.[c,16] These experiments were presented as an illustration of how a non-expert could be equipped with the know-how to develop a bioweapon.[17] This demonstrated that LLMs such as ChatGPT can provide users with information that is already publicly available, slightly lowering the barriers to bioweapons production.[d,18]

b The availability of information on genetic material that could potentially create harm or be misused is increased. Individuals who do not have peaceful or bonafide reasons can order or synthesise oligonucleotides with infection-causing abilities, thereby causing harm to the population.

c Some DNA companies have voluntarily signed up with the International Gene Synthesis Consortium (an industry-led group) that accounts for gene synthesis protocols to mitigate abuse of the technology.

d Although this study received attention in the media, a biodefence expert opined that it was a poorly designed study. These insights were revealed through an email communication with the author.

Biological Design Tools (BDTs) form another group of AI tools used in the life sciences. BDTs are trained on biological sequence data and aid in the design and engineering of biological molecules, processes, or systems.[19] BDTs, such as AlphaFold2 and MegaSyn, help synthesise novel biomolecules for beneficial purposes, such as for pharmaceutical purposes or in the design of vaccine candidates.[20] BDTs are prone to risks as they can contribute to the development of biological molecules or pathogens with selective advantages, including enhanced viral infectivity or resistance to existing antibiotic therapies.[21] Another likely scenario involves the use of AI-bio tools to direct a slight modification in an existing pathogen to make it more infectious to a certain population.[22] In addition, possibilities exist wherein BDTs could help malevolent actors conceal detection through sequencing-based screening, making attribution[e] difficult.[f,23]

e Attribution refers to the identification of the origin of a biological incident that is carried out to determine whether the incident was "natural, deliberate, or accidental" (See: https://centerforhealthsecurity.org/sites/default/files/2023-02/20230124-bioattribution-mtg-rpt.pdf). The concealment of the origin of a biological incident by AI-bio tools would be counterproductive to any forensic investigation, hamper accountability, weaken biodefence capabilities, and impede dissuasion policies against bioweapons.

f AI tools are increasingly being used to aid in attribution or in identifying the origin of a DNA sequence. However, the possibility exists that advances in AI-bio could lead to situations where AI enables evasion of attribution.

Rationale for Bioweapons Threat

AI remains a "data-based system",[24] and its capabilities are limited by the nature and amount of biological information that is incorporated into it and the intent of its users. Further, there is a gap between the digital and the physical space. The digital-physical divide refers to the junction at which the design for a biological process or molecule moves from a digital design to the physical space or to the production of a bioproduct. An AI tool could provide a detailed workplan of how to create a pathogen, but detailed experimental work would prove to be a challenge and a rate-limiting step in the design being converted into an actual pathogen.[25]

Bridging this divide relies on two primary factors: scientific training and intent.[g,26] Researchers have stressed that technical laboratory experience, tacit knowledge, and troubleshooting are essential to the success of any biological experiment.[27] Current BDTs are "limited by users' ability to express what they want in a language that the models can interpret".[28] For instance, if a user wants to enhance the surface properties of a biological molecule, then they would have to input specific parameters based on their technical expertise into the tool.[29] The current technological capabilities of AI can only facilitate knowledge transfer and may not be able to address everyday issues that arise from biological experiments that require hands-on experience acquired tacitly and through trial and error. Moreover, LLMs often "hallucinate", which can cause users who are not scientifically trained to make errors that adversely impact the fundamental working of experiments.[h]

Additionally, biological experiments require expensive, specialised equipment and materials, which may only be available in specific laboratories. Moreover, the availability of high-quality and complete biological datasets, such as DNA or protein sequences, to train AI is uncertain as this depends upon proprietary rights, licensing policies, ethical, and security considerations.[30] Incomplete datasets are problematic in experimental design, giving rise to variable outcomes, and can be impediments to bioweapons development.[31] Collectively, these factors have been cited as significant barriers to bioweapons development by various non-state actors and have been responsible for failed attempts.[i,32]

---

g Unlike nuclear weapons—where the barriers to its development depend on the acquisition of specific material, such as highly enriched uranium and plutonium—the development of bioweapons relies on materials that can be used for beneficial purposes and for creating harm, and are thus 'dual-use' in nature. In addition, bioweapons development relies heavily on trial-and-error and experimental validation processes, which come about through scientific training and experience.

h If a non-expert were to receive "hallucinations", then the non-expert would be unable to identify whether the information was inaccurate.

i Aum Shinrikyo and al-Qaeda's attempts were impeded by socio-technical and organisational challenges. In addition, policymakers claim that bioweapons development is impeded by technical challenges, including reproducibility of experiments, conversion to large-scale weapons deployment, and weaponisation of existing pathogens. See: https://www.tandfonline.com/doi/full/10.1080/0163660X.2020.1770969.

Rationale for Bioweapons Threat

The inclusion of AI has not reduced the barriers for converting technical knowledge to a biological agent; instead, it has made the acquisition of technical information easier and faster.

It is imperative to understand the threat landscape in a contextual manner and this requires an understanding of a country's national security landscape. At an international level, the use of biological weapons, with or without AI, can be carried out by rogue states or through non-state actors.[33] Typically, terror outfits partake in activities that induce fear and panic in the state and the population and are driven by a particular cause.[34] In India, the greatest threats are posed by ethnic nationalists and separatists who want broad political support or international recognition and are unlikely to use weapons that violate norms/taboos.[35] The use of a biological weapon is thus an unlikely weapon of choice as it may harm the outfit's support base.[j,36] Lone-wolf attacks, such as by a disgruntled scientist with access to technology and resources or state-sponsored attacks, similar to the alleged Chinese cyber-espionage threats against India, remain unlikely occurrences.[37] The risk of being apprehended is extremely high to warrant any benefits from a clandestine bioweapons attack. Additionally, such attacks require sustained access to extremely sophisticated equipment and specialised materials.[38]

It is increasingly evident, however, that AI may be useful in identifying vulnerabilities in existing critical biological infrastructures. For instance, analysts from the cyberbiosecurity[k] domain posit that AI may manipulate or steal biological data or interfere with experimental settings. For instance, an AI-bio tool could be used to interfere with the temperature settings on refrigerated RNA samples, thus compromising their biological integrity.[39] Another concern is that LLMs could direct malevolent actors to target vulnerable agricultural ecosystems and disrupt the food chain.[40] Additionally, a data-poisoning attack[l] can skew or "poison scientific knowledge".[41] In addition, LLM hallucinations create opportunities for misinformation, hampering biosecurity efforts during public health emergencies, and promote opportunities for disinformation through fake news.[42] Thus, systematic threat assessments on the risk of a biological attack would provide detailed information and ensure an appropriate level of preparedness.

---

j The employment of bioweapons by state actors is unlikely because the Biological Weapons Convention (BWC) has 109 signatories. In the case of non-state actors, intent depends upon the cause of the malevolent actor.

k A field at the intersection of cybersecurity and biological data. Cyberbiosecurity is becoming increasingly relevant owing to the large amounts of biological data that is stored digitally and the value of information it stores.

l The manipulation of the data that an AI model is trained on.

Implications of the AI-Bio technology Convergence

Through stakeholder discussions and an analysis of the existing literature, it is evident that there is considerable publicity around the intersection of AI with biotechnology and its potential for misuse.[43] AI by itself is not a technology, but it can make existing technologies work more efficiently.[44] Advancements in AI are progressing rapidly, necessitating a biosecurity framework that addresses the risks posed by the convergence of AI with biotechnology. Scientists in the UK and the US have voiced grave concerns regarding its risks, which culminated in the Bletchley Declaration.[m,45]

In the context of India, NITI Aayog addressed the ethical aspects of AI in its 2022 "Responsible AI For All" document, while the Indian Council of Medical Research (ICMR) released its ethical guidelines on the use of AI in biomedical research and in healthcare in 2023.[46] Earlier this year, India's External Affairs Minister S. Jaishankar warned against the potential dangers of AI in a globalised world.[47] However, current policy discussions in India do not account for the potential misuse of AI-bio tools for bioweapons production, which has resulted in this issue being relegated to a policy vacuum. Moreover, while India has biosafety frameworks, particularly for the conservation of biodiversity and the responsible use of biological materials, it lacks a comprehensive biosecurity framework that encompasses bioweapons and biowarfare.[48]

Policymakers need to formulate a dissuasion policy aimed at discouraging malevolent actors from bioweapons development using AI tools. This entails the adoption of a set of "actions taken to increase the target's perception of the anticipated costs and/or decrease its perception of the likely benefits from developing, expanding, or transferring" a capability "that would be threatening".[49] Highlighting the difficulties associated with creating a bioweapon using AI and dispelling popularised notions of the ease with which AI-bio tools can aid in bioweapons development can create a strong barrier to entry. Thus, a dissuasion policy would discourage a malevolent actor from initiating or expanding a capability on the assumption that the "anticipated costs significantly outweigh the benefits".[50]

m A global agreement on the safe design, development, and deployment of AI, of which India is a signatory.

In addition, a framework would enable India to assess threats that arise at a transnational level; for instance, if an open-source BDT developed in India were to be misused by a foreign actor, then India would have the capability to address the threat.[51] The development of such a policy relies on understanding the bioweapons threat landscape in India, the available AI-bio tools, and accessible biological datasets. Therefore, its application in the Indian context requires a nuanced approach.[n,52]

n The Bletchley Declaration is the first global pact that recognises the potential benefits of AI as well as the risks posed by it.

India needs to address the risks posed by the AI-bio convergence through a biosecurity framework, where the potential misuse of biotechnology by AI can be elucidated by involving international and national stakeholders from the disarmament community, scientists from the life sciences, government officials, experts from industry and academia, the intelligence community, and civil society.[o,53]

Formulation of Threat Assessments and Evaluation of AI Models

As with all emerging threats, there is a need to inculcate situational awareness and detection systems within frameworks to detect biological attacks.[54] Appropriate threat assessments can be formulated through red-teaming strategies.[p,55] A recent red-teaming exercise by RAND Corporation found that malevolent actors would find it difficult to develop bioweapons by using existing AI tools but cautioned against the risks posed by future AI tools and stressed on the need for regular red-teaming exercises.[56] India can participate in similar exercises to develop an understanding of the capabilities of AI in bioweapons development in the Indian context. In addition, threat assessments need to incorporate other contributing factors to bioweapons development, such as the dark web, the use of unmanned aerial vehicles (drones), and 3D printing, which are known to aid in access to materials and equipment and the delivery of biological weapons.[57]

Implementation of Nucleic Acid Synthesis Screening Practices

The International Gene Synthesis Consortium (IGSC), International Biosafety and Biosecurity Initiative for Science (IBBIS), and SecureDNA are groups that employ a voluntary system for companies to determine the legitimacy of nucleic acid sequences that are ordered to prevent its misuse either deliberately or accidentally.[58] This involves screening nucleic acids to identify and track ones that are of concern. Earlier this year, the US made it mandatory for government-

o Any policy framework to address the risks associated with emerging technologies needs to take into consideration interactions with the intelligence community. There is a need to convey to the intelligence community nuances about bioweapons development and how to detect potential threats, which would ultimately enable future policies to work better. This also requires extensive collaboration with biosecurity experts.

p A mixed-methods approach was used to ascertain the security of a system and to enhance biosecurity. This constituted a red-teaming exercise, where experts emulated malevolent actors trying to develop biological weapons. This was used to determine various risk scenarios and offer insights into policy and regulatory frameworks governing AI and biotechnology.

funded projects to acquire nucleic acids from companies that adhere to best practices in nucleic acid screening.[59] This step is aimed at driving companies that want to fund US federal government projects to mandatorily screen orders. India can implement a similar know-your-customer (KYC) based approach in the acquisition of biological materials and make it mandatory for companies to adopt best practices in screening for nucleic acids.[q]

Inclusion of Technical Barriers to AI Models

Effective guardrails to prevent the AI-bio malign risk could include engaging with AI tool developers to ensure that appropriate biological data is available to the tools or restricting the use of certain datasets.[60] This would aid developers in assessing potentially dangerous activities and flagging them to be addressed and would act as a regulatory mechanism to account for potential threats.[61]

Refusals, or when an AI model refuses to adhere to user requests that are construed as alarming, could be incorporated as another safeguard to prevent harmful actions.[62] In addition, educating and incentivising research organisations, such as those specialising in synthetic biology and AI tool developers, on potential areas of misuse would be a useful bottom-up strategy to inculcate a sense of responsibility and accountability over these disruptive technologies.[63]

Induction of AI Safety Institutes

AI Safety Institutes as governance structures have been set up in some countries to ensure that AI receives adequate infrastructural support to harness the technology while gaining an understanding of its risks.[64] Current policy discourses are divided around setting up AI Safety Institutes in India; however, the AI Safety Institute has been proposed to be one that hosts a "series of connections to innovation-led and existing community-driven ecosystems, focused on different aspects of AI."[65] As a signatory to the Bletchley Declaration, India can implement AI Safety Institutes to ensure maximal innovation through AI while addressing the risks that they can cause by placing biosecurity as a thematic domain of concern. Moreover, safety is a prominent pillar in the IndiaAI Mission policy, which has a budget of INR 10,000 crores.[r,66]

q Despite ongoing discussions, the BWC does not have a verification and compliance mechanism. An AI-bio framework could pave the way for promoting regulatory oversight in bioweapons development.

r The IndiaAI Mission policy aims to harness the potential benefits of AI in various sectors and promote an ecosystem that fosters innovation in AI. The policy also aims to democratise computing, develop indigenous models, and promote the development of AI technologies that are ethical.

Conclusion

The ease with which LLMs can be used to help create pathogens has alarmed the policy community. Although current AI-bio capabilities do not pose a significant risk to the development of bioweapons, AI-bio tools are capable of slightly lowering the barriers to bioweapons development by presenting users with information that is publicly available in an efficient manner. Nevertheless, there exists a significant barrier in translating information that is available in the digital space to the physical one.

The current discourse on bioweapons development with aid from AI-bio tools is being magnified by media reports. This underscores the importance of developing nuanced understandings of disruptive technologies, which can be achieved through stakeholder discussions with the biotechnology community and AI developers. The over-estimation of AI-bio tools can hamper dissuasion policies. There is an urgent need to balance the publicity from media reports because it can persuade malicious actors to undertake development based on the perception that AI has simplified the production of bioweapons.

AI-bio tools may present severe threats in the future. The rapid pace of development of these technologies and their potential for misuse by malevolent actors necessitates the formulation of a biosecurity framework that incorporates AI. Effective guardrails would ensure that appropriate mechanisms are installed to detect and respond to AI-aided biological threats. Such frameworks would deter the misuse of biological research for weaponisation and deter malicious actors from bioweapons development activities.

Lakshmy Ramakrishnan *is Associate Fellow with the Health Initiative, Observer Research Foundation.*

The author wishes to acknowledge the expert comments made by the following individuals on an early draft of this brief: Group Captain (Dr) Ajey Lele (Rtd.), Manohar Parrikar Institute for Defense Studies and Analysis, India; Dr Suryesh K. Namdeo, Department of Science and Technology – Centre for Policy Research, India; Dr Gregory D. Koblentz, George Mason University, US; and Dr Pawan K. Dhar, CVJ Centre for Synthetic Biology and Biomanufacturing, India.

Endnotes

1 Trond Arne Undheim, "The Whack-a-Mole Governance Challenge for AI-Enabled Synthetic Biology: Literature Review and Emerging Frameworks," *Frontiers in Bioengineering and Biotechnology* 12 (2024), https://doi.org/10.3389/fbioe.2024.1359768.

2 "Governing AI For Humanity," United Nations, September 2024, https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

3 Isaias Ghebrehiwet, Nazar Zaki, Rafat Damseh, and Mohd Saberi Mohamad, "Revolutionizing Personalized Medicine with Generative AI: A Systematic Review," *Artificial Intelligence Review* 57, no. 5 (2024), https://doi.org/10.1007/s10462-024-10768-5.

4 "Governing AI for Humanity"

5 "Governing AI for Humanity"

6 "Governing AI for Humanity"

7 "A.I. May Save Us or May Construct Viruses to Kill Us," *The New York Times*, July 2024, https://www.nytimes.com/2024/07/27/opinion/ai-advances-risks.html; "Could Chatbots Help Devise the next Pandemic Virus?" *Science*, June 2023, https://www.science.org/content/article/could-chatbots-help-devise-next-pandemic-virus; Steph Batalis, "Can Chatbots Help You Build a Bioweapon?" *Foreign Policy* (blog), December 12, 2024, https://foreignpolicy.com/2023/11/05/ai-artificial-intelligence-chatbot-bioweapon-virus-bacteria-genetic-engineering/

8 "AI-Made Bioweapons Are Washington's Latest Security Obsession," *Bloomberg*, August 2, 2024, https://www.bloomberg.com/news/features/2024-08-02/national-security-threat-from-ai-made-bioweapons-grips-us-government; The White House, "FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence," October 2023, 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/; "Prime Minister's Speech at the AI Safety Summit," UK Government, November 2023, https://www.gov.uk/government/speeches/prime-ministers-speech-at-the-ai-safety-summit-2-november-2023

9 "Red-Teaming the Risks of Using AI in Biological Attacks," RAND, March 25, 2024, https://www.rand.org/pubs/articles/2024/red-teaming-the-risks-of-using-ai-in-biological-attacks.html

10 Matt Field, "Apathy and Hyperbole Cloud the Real Risks of AI Bioweapons," *Bulletin of the Atomic Scientists*, September 12, 2024, https://thebulletin.org/2024/09/apathy-and-hyperbole-cloud-the-real-risks-of-ai-bioweapons/

11 Milton Leitenberg, "Assessing the Biological Weapons and Bioterrorism Threat ," December 2005. https://press.armywarcollege.edu/cgi/viewcontent.cgi?article=1029&context=monographs; C. Caverly, ""When the Enemy Drew Our Attention": Reconsidering Prior "When the Enemy Drew Our Attention": Reconsidering Prior Restraint in the Context of Dual Use Research of Concern Restraint in the Context of Dual Use Research of Concern," *William and Mary Bill of Rights* 28, no. 1 (2019).

Endnotes

12 OpenAI, "GPT-4 System Card," February 23, 2023, https://cdn.openai.com/papers/gpt-4-system-card.pdf.

13 OpenAI, "GPT-4 System Card"

14 OpenAI, "GPT-4 System Card"

15 "Could Chatbots Help Devise the Next Pandemic Virus?"

16 "Could Chatbots Help Devise the Next Pandemic Virus?"

17 "Could Chatbots Help Devise the Next Pandemic Virus?"

18 Field, "Apathy and Hyperbole Cloud the Real Risks of AI Bioweapons"; Insights revealed to the author during an email communication with a biodefence consultant.

19 "The Convergence of Artificial Intelligence and the Life Sciences," Nuclear Threat Initiative, October 2023, https://www.nti.org/analysis/articles/the-convergence-of-artificial-intelligence-and-the-life-sciences/

20 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

21 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

22 Insights revealed to the author during a personal communication with a synthetic biologist.

23 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

24 Insights revealed to the author during an interview with a CBRN policy expert.

25 Insights revealed to the author during an interview with a biosecurity consultant.

26 Insights revealed to the author during an interview with a CBRN policy expert.

27 Field, "Apathy and Hyperbole Cloud the Real Risks of AI Bioweapons"

28 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

29 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

30 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

31 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

Endnotes

32 Gregory D. Koblentz, "Emerging Technologies and the Future of CBRN Terrorism," *The Washington Quarterly* 43, no. 2 (2020): 177–96, doi:10.1080/0163660X.2020.1770969.

33 Insights revealed to the author during an interview with a CBRN policy expert.

34 Insights revealed to the author during an interview with a CBRN policy expert.

35 Insights revealed to the author during a personal email communication with a bioweapons expert.

36 Insights revealed to the author during a personal email communication with a bioweapons expert.

37 Insights revealed to the author during an interview with a biosecurity consultant.

38 Insights revealed to the author during an email communication with a biodefence consultant.

39 Insights revealed to the author during an interview with a biosecurity consultant.

40 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

41 J. Yang et al., "Poisoning Scientific Knowledge Using Large Language Models," bioRxiv, November 10, 2023, https://doi.org/10.1101/2023.11.06.565928.

42 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

43 Insights revealed to the author during interviews with a CBRN policy expert and with a biosecurity consultant.

44 Insights revealed to the author during an interview with a CBRN policy expert.

45 "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023," UK Government, November 2023, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

46 "Ethical Guidelines for Application of Artificial Intelligence in Biomedical Research and Healthcare," Indian Council of Medical Research, 2023, https://main.icmr.nic.in/sites/default/files/upload_documents/Ethical_Guidelines_AI_Healthcare_2023.pdf; "Responsible AI," NITI Aayog, November 2022, https://www.niti.gov.in/sites/default/files/2022-11/Ai_for_All_2022_02112022_0.pdf?trk=public_post_comment-text.

47 "Jaishankar Cautions "AI Is Just as Dangerous as Nuclear Weapons". Here's Why," *LiveMint*, October 6, 2024, https://www.livemint.com/ai/artificial-intelligence/jaishankar-cautions-ai-is-just-as-dangerous-as-nuclear-weapons-heres-why-11728210010601.html.

48 Shravishtha Ajaykumar, "Biosecurity Blueprint for India," Observer Research Foundation, August 2024, https://www.orfonline.org/expert-speak/biosecurity-blueprint-for-india.

Endnotes

49 "Dissuasion Strategy," Centre for Strategic and Budgetary Assessments, 2008, https://csbaonline.org/uploads/documents/2008.05.06-Dissuasion-Strategy.pdf.

50 "Dissuasion Strategy"

51 Insights revealed to the author during an email communication with a biodefense consultant.

52 "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023," UK Government, November 2023, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

53 Shravishtha Ajaykumar, "Global Collaboration in Integrated AI and Biotech," Observer Research Foundation, December 2023, https://www.orfonline.org/expert-speak/global-collaboration-in-integrated-ai-and-biotech.

54 Insights revealed to the author during an interview with a CBRN policy expert.

55 Lisa Zhang and Gigi Kwik Gronvall, "Red Teaming the Biological Sciences for Deliberate Threats." *Terrorism and Political Violence* 32, no. 6 (2018): 1225–44, doi:10.1080/09546553.2018.1457527.

56 "Red-Teaming the Risks of Using AI in Biological Attacks"

57 Koblentz, "Emerging Technologies and the Future of CBRN Terrorism"

58 "Biosecurity Innovation and Risk Reduction: A Global Framework for Accessible, Safe and Secure DNA Synthesis," Nuclear Threat Initiative , January 2020, https://media.nti.org/documents/Biosecurity_Innovation_and_Risk_Reduction.pdf.

59 "Framework for Nucleic Acid Synthesis Screening," The White House, April 2024, https://www.whitehouse.gov/ostp/news-updates/2024/04/29/framework-for-nucleic-acid-synthesis-screening/.

60 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

61 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

62 Nuclear Threat Initiative, "The Convergence of Artificial Intelligence and the Life Sciences"

63 "Sounding the Alarm on AI-Enhanced Bioweapons," European Leadership Network, February 2024, https://europeanleadershipnetwork.org/commentary/sounding-the-alarm-on-ai-enhanced-bioweapons/.

64 "Introducing the AI Safety Institute," UK Government, January 2024, https://www.gov.uk/government/publications/ai-safety-institute-overview/introducing-the-ai-safety-institute.

Endnotes

65 Carnegie Endowment for International Peace, "Disrupting AI Safety Institutes: The India Way," https://carnegieendowment.org/posts/2024/09/disrupting-ai-safety-institutes-the-india-way?lang=en.

66 "Cabinet Approves Ambitious IndiaAI Mission to Strengthen the AI Innovation Ecosystem," March 2024, https://pib.gov.in/pib.gov.in/Pressreleaseshare.aspx?PRID=2012355.

Images used in this paper are from Getty Images/Busà Photography.

ORF OBSERVER RESEARCH FOUNDATION

Ideas . Forums . Leadership . Impact

20, Rouse Avenue Institutional Area,
New Delhi - 110 002, INDIA
Ph. : +91-11-35332000. **Fax :** +91-11-35332005
E-mail: contactus@orfonline.org
Website: www.orfonline.org