HSP
HARVARD SUSSEX PROGRAM

# Artificial Intelligence Technologies and Chemical and Biological Weapons: A Chronology of Events (2000-Present)

Joshua Moon

Boaz Chan

Alexander Ghionis

Shaunna McIvor

Working version

# INTRODUCTION

This chronology presents a working timeline of events and policy documentation of relevance to the theme of Artificial Intelligence in the Chemical and Biological Weapons Prohibition Regimes. This chronology is a work in progress and thus does not represent a 'final' product. Our intention is to continue updating this chronology as new events occur. All authors have been involved in identifying events, adding entries and providing sources.

Entries in the chronology were chosen according to the following parameters:

- Must focus on AI technology
- Must be concerned with CBW/CB Security
- May include general LAWS/ WMD/ CBRNe/ Public Health relevance

Where possible, all entries in this chronology are based on publicly available and open-access sources. This has enabled us to archive our sources in a separate folder, providing validation for the chronology's timeline. Not all of our source materials, however, are easily downloadable documents or widely available on CC-BY license. For example, embedded videos on websites or HTML text that isn't readily downloadable. Second, not all our source materials are free of copyright, and we therefore cannot capture these to share along with the chronology. In either of these instances, the citation below the entry will include hyperlinks to the relevant pages. Source material(s) can be made available upon request.

If you wish to contact the team please email:

Dr Joshua Moon – j.r.moon@sussex.ac.uk

Dr Alexander Ghionis – a.ghionis@sussex.ac.uk

# Contents

# CHRONOLOGY ENTRIES

## 2015-2021

**30 November 2015.** Jonathan Forman, the OPCW's Science Policy Adviser, hosts a side-event at the 20th Session of the Conference of the States Parties (held at the World Forum in The Hague, the Netherlands) titled "Emerging Technologies and the CWC: Autonomous Systems and Artificial Intelligence"[1]

**[UI] 11 February 2019.** The United States Department of Defense (DoD) launches its first artificial intelligence (AI) strategy serve as a guideline on the principles that they plan on developing related technologies. The overall theme of the DoD's strategy report revolves around three core principles: fostering innovation, maintaining national security, and upholding ethical and responsible AI development and use. The DoD recognizes the transformative impact of AI on military capabilities and has placed an emphasis on promoting innovation and technological advancement. It actively invests in AI research and development, collaborates with the private sector, and fosters partnerships with academia to stay at the forefront of AI technology.[2]

**10 - 14 June 2019.** The OPCW's Scientific Advisory Board's 28th Session featured agenda items in relation to artificial intelligence. In particular, Subitem 11(a) Artificial intelligence (AI) for chemical verification; Subitem 13(a): Digital transformation powered by AI and related cybersecurity considerations; and, Subitem 13(b): Digitalisation in the chemical industry.[3]

**20 October 2020.** The European Parliament's "Framework of ethical aspects of artificial intelligence, robotics and related technologies" [1] stresses, inter alia,

> "the importance of the creation of an ethical code of conduct underpinning the deployment of weaponised AI-enabled systems in military operations, similar to the existing regulatory framework prohibiting the deployment of chemical and biological weapons" [CBW comparison]
>
> "is of the opinion that the Commission should initiate the creation of standards on the use of AI-enabled weapons systems in warfare in accordance with international humanitarian law, and that the Union should pursue the international adoption of such standards" [EC as standard-setter]
>
> "considers that the Union should engage in AI diplomacy in international fora with like-minded partners like the G7, the G20 and the OECD"[4]

**13 May 2021.** The UK Ministry of Defence and Germany's Bundeswehr Office for Defence Planning jointly publish "Human Augmentation – The Dawn of a New Paradigm", a strategic implications project report examining the future of human augmentation technologies and their potential impacts on society and defense. While the report does not focus specifically on artificial intelligence in relation to chemical and biological weapons, it does

---

[1] Forman, J. (2015) *Autonomous Systems and AI*. Available at: https://www.opcw.org/sites/default/files/documents/Science_Technology/Diplomats_Programme/20151130-Autonomous_systems_and_AI-JForman-Print.pdf (Accessed: 4 October 2024)

[2] Cronk, T. (2019) DOD Unveils Its Artificial Intelligence Strategy. Available at: https://www.defense.gov/News/News-Stories/Article/Article/1755942/ (Accessed: 4 October 2024)

[3] OPCW Scientific Advisory Board. (2019) *Report of the Scientific Advisory Board at its Twenty-Eighth Session, 10–14 June 2019*. Available at: https://www.opcw.org/sites/default/files/documents/2019/09/sab-28-01%28e%29_0.pdf (Accessed: 4 October 2024

[4] European Parliament. (2020) *European Parliament Resolution of 20 October 2020 with Recommendations to the Commission on a Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies*. Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html (Accessed: 4 October 2024)

discuss AI as an enabling technology for human augmentation that could have implications for warfare and weapons development.

The report frames AI as a key component of future human-machine teaming, stating "The winners of future wars will not be those with the most advanced technology, but those who can most effectively integrate the capabilities of people and machines." It notes that AI could help address cognitive overload issues as warfare becomes more complex and data-intensive.

In terms of potential risks, the report briefly mentions that "Synthetic biology could be used to rapidly develop new treatments and vaccines. It could also be used for producing biological weapons such as designer viruses and diseases." However, it does not explicitly link this to AI capabilities.

The authors emphasize the need for early engagement on the ethical and legal implications of human augmentation technologies, including AI. They note that "Defence cannot wait for ethics to change before engaging with human augmentation, we must be in the conversation from the outset to inform the debate and understand how ethical views are evolving."

Overall, while not focused on CBW specifically, the report highlights AI as a transformative technology that will shape future warfare capabilities and presents both opportunities and risks that defense organizations must proactively address. The framing suggests AI will be a critical enabler of enhanced human performance and human-machine teaming, rather than an autonomous weapons technology.[5]

**15 December 2021.** The European Parliament's resolution on "the challenges and prospects for multilateral weapons of mass destruction arms control and disarmament regimes ( )" refers to AI in the following terms[6]

- "having regard to its resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies"[7] [see Chronology entry for 20 October 2020]
- 2. "[...] Is alarmed by the ongoing erosion of the global non-proliferation, disarmament and arms control architecture, which is worsened by the rapid development of new potentially destabilising systems, such as weapon systems enabled with artificial intelligence (AI)" [is the narrative continued or shifted? How?]
- 36. "[...] call for the EU to also pave the way for global negotiations to update all existing arms control, disarmament and non-proliferation instruments, so as to take AI-enabled systems used in warfare into account" [Update existing arms control is useful, what are they doing so far?]

## MARCH 2022

**7 March 2022.** Fabio Urbina, Filippa Lentzos, Cédric Invernizzi and Sean Ekins publish a commentary in Nature Machine Intelligence that reports on an experiment in which a commercial de novo molecule generator (known as MegaSyn) was provided with thresholds that drove it toward generating compounds similar to VX. They report it took 6 hours to generate 40,000 molecules, including VX, and "many other known chemical warfare agents [...] Many new molecules were also designed that looked equally plausible." They suggest that it is "entirely possible that novel routes can be predicted for chemical warfare agents, circumventing national and international lists of watched or controlled precursor chemicals for known synthesis routes" and that their experiment demonstrates

[5] UK Ministry of Defence 'Human Augmentation – The Dawn of a New Paradigm' HMG 13 May 2021 available at https://www.gov.uk/government/publications/human-augmentation-the-dawn-of-a-new-paradigm (accessed 30/09/2024)
[6] European Parliament resolution of 15 December 2021 on the challenges and prospects for multilateral weapons of mass destruction arms control and disarmament regimes (2020/2001(INI)), available at: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0504_EN.html
[7] European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)) – see Chronology entry for 20 October 2020

"the potential for dual-use" with such AI systems.[8] This commentary was widely cited by other journal articles and mainstream media sources filled with recommendations and concern for the restriction of AI in the development of chemical/biological weapons as it acted as a wakeup call to its nefarious potential for synthesizing chemical and biological warfare agents[9]

## JUNE 2022

**16-17 June 2022.** The OPCW's Scientific Advisory Board and the International Union of Pure and Applied Chemistry (IUPAC) jointly organised an international workshop on 'Artificial Intelligence Assisted Chemistry', held in the Hague, Netherlands.[10] Speakers include: Prof. Connor Coley, Massachusetts Institute of Technology (MIT, USA); Prof. Richard Bourne, University of Leeds (UK); Prof. Yuan Yao, Yale University (USA); Dr. Marwin Segler, Microsoft; Prof. Jeremy Frey, University of Southampton (UK); Dr. Sean Ekins, Collaboration Pharmaceuticals Inc. (USA); and Dr. B. Saha, Dean, Sagar Group of Institutions, India.[11] Each speaker spoke on a specific topic of which they were experts on themselves such as Prof. Jeremy Frey, who spoke on the topic of the need for "smart and intelligent" labs when integrating machine learning with chemistry. Prof. Frey recently published a journal article on responsible research and innovation of AI in the food industry.[12]

## JULY 2022

**27-28 July 2022.** The U.S. Defense Department plans to invest an additional $300 million per year over the next five years to safeguard against known and emerging biological threats. At the National Defense Industrial Association's Chemical, Biological, Radiological, and Nuclear Defense Exhibition and Conference, Deb Rosenblum, Assistant Secretary of Defense for nuclear, chemical, and biological programs, highlighted the growing concept of "bio-convergence," which involves the fusion of biological sciences with emerging technologies, including artificial intelligence (AI), nanotechnology, and physics.[13] This increase in spending will fund initiatives to bolster its defense capabilities against chemical, biological, radiological, and nuclear (CBRN) threats. These measures include equipping various military platforms with CBRN sensors, enhancing satellite and thermal imaging with advanced algorithms, and developing modern vaccines that can both protect and treat against CBRN agents. Notably, the focus extends to individual soldiers, who may soon have access to wearable devices capable of providing real-time updates on CBRN exposure. Alongside these technological advancements, a new strategy emphasizes the importance of integrating a CBRN defense mindset and culture

---

[8] Urbina, F., Lentzos, F., Invernizzi, C. and Ekins, S. 'Dual use of artificial-intelligence-powered drug discovery' Comment (Nature Machine Intelligence, 2002) 4

[9] See both 'Much to discuss in AI ethics' Nat Mach Intell 4, 1055–1056 (2022). https://doi.org/10.1038/s42256-022-00598-x and The Economist. (2022, March 19). How to Tweak Drug-Design Software to Create Chemical Weapons. Retrieved from https://www.economist.com/science-and-technology/how-to-tweak-drug-design-software-to-create-chemical-weapons/21808200

[10] For a full report see Saeed, A., Hotchkiss, P. et al. ' Artificial Intelligence-Assisted Chemistry' (*Chemistry International,* 2023) July-September, available at https://www.degruyter.com/document/doi/10.1515/ci-2023-0314/html (accessed 19/07/2024)

[11] '**International Union of Pure and Applied Chemistry (IUPAC).** (2022) *International Workshop on Artificial Intelligence Assisted Chemistry*. Available at: https://iupac.org/wp-content/uploads/2022/07/International_workshop_on_Artificial_Intelligence.pdf (Accessed: 4 October 2024)

[12] Craigon, Peter J., Sacks, Justin, Brewer, Steve, Frey, Jeremy, Gutierrez, Anabel, Jacobs, Naomi, Kanza, Samantha, Manning, Louise, Munday, Samuel, Wintour, Alexsis and Pearson, Simon (2023) Ethics by design: responsible research & innovation for AI in the food sector. Journal of Responsible Technology, 13, 2666-6596

[13] P **Magnuson, S.** (2022) *JUST IN: Pentagon Biological Defense Programs at 'Pivot Point'*. Available at: https://www.nationaldefensemagazine.org/articles/2022/7/28/pentagon-biological-defense-programs-at-pivot-point (Accessed: 4 October 2024).

within the military. These combined efforts aim to ensure more effective detection, response, and protection against CBRN threats while reducing unnecessary battlefield restrictions.[14]

## SEPTEMBER 2022

**1-14 September 2022.** The fifth Spiez Convergence conference is held on 1-2 (ice-breaker event) and 11-14 September, in person at the Spiez Laborary, Switzerland. In the Executive Summary of the Conference Report, it was noted that:

> "One year ago, during Spiez CONVERGENCE 2021, a presentation demonstrated the power of AI for discovering new toxic chemicals. The resulting publication "Dual use of artificial-intelligence-powered drug discovery" had a strong media impact worldwide. Subsequently, a second publication "A teachable moment for dual-use" discussed more broadly the implication for the AI community as well as for the scientific community. The technologies for Machine Learning and Artificial Intelligence are close to becoming Game Changers; they may profoundly affect the regimes prohibiting chemical and biological weapons. The combination of AI with synthetic biology, automation and robotics, Big Data, high-throughput synthesis and screening, leads to a context shift in how experiments are performed." *(p.10)*[15]

**12 September 2022.** The Executive Order on the Bioeconomy is a comprehensive plan to advance the United States' bioeconomy, with a primary focus on biotechnology and biomanufacturing. This order places a strong emphasis on data-driven initiatives, aiming to identify crucial data types and sources while addressing data gaps, enhancing security, and mitigating privacy risks associated with these technologies. It also seeks to bolster domestic biomanufacturing capacity, improve processes, and ensure biosafety and biosecurity, while addressing risks from foreign adversaries. The order promotes the procurement of biobased products within federal agencies and encourages workforce development in biotechnology and biomanufacturing, particularly with a focus on equity and underserved communities. Regulatory clarity and efficiency are a key component, with efforts to simplify and streamline regulations for the biotechnology sector. It also introduces an initiative to reduce biological risks and measures the economic value of the bioeconomy. Lastly, the order directs a comprehensive assessment of national security threats associated with foreign adversaries' involvement in the bioeconomy and encourages international collaboration in research, regulatory practices, and data sharing to support both the U.S. and global bioeconomies. In essence, this executive order aims to harness the potential of biotechnology for innovation and economic growth while addressing significant challenges and mitigating risks.[16]

## JANUARY 2023

**13 January 2023.** UNICRI published its Handbook to combat CBRN disinformation. The press release notes that "Chemical, biological, radiological, and nuclear (CBRN) disinformation is defined as intentionally misleading and deceptive information about CBRN threats that can potentially cause serious political, financial, and physical harm to governments, international organizations, the scientific community, academia, industry, and

---

[14] **South, T.** (2022) *Big Changes Ahead for How Troops Battle Future Chemical, Biological Threats*. Available at: https://www.militarytimes.com/news/2022/08/02/big-changes-ahead-for-how-troops-battle-future-chemical-biological-threats/ (Accessed: 4 October 2024)

[15] **Moodie, A. and Revill, J.** (2022) *Spiez Convergence: Report on the Fifth Conference*. Available at: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/582482/SpiezConvergenceReport-2022.pdf?sequence=1 (Accessed: 4 October 2024)

[16] **The White House.** (2022) *Executive Order on Advancing Biotechnology and Biomanufacturing Innovation for a Sustainable, Safe, and Secure American Bioeconomy*. Available at: https://www.whitehouse.gov/briefing-room/presidential-actions/2022/09/12/executive-order-on-advancing-biotechnology-and-biomanufacturing-innovation-for-a-sustainable-safe-and-secure-american-bioeconomy/ (Accessed: 4 October 2024).

the population at large."[17] References to AI tend to be implicitly included in discussion about platforms and technologies, although one explicit reference from page 32 says:

> "A deepfake video is the product of an Artificial Intelligence (AI) technique for human image synthesis that combines and superimposes existing images and videos onto source images or videos. These videos or photographs can misrepresent people by generating images that are nearly indistinguishable from the original. If combined with speech synthesis systems (that learn to imitate individuals' voices), deepfake videos can misrepresent people by reproducing not only their voices, but also their cadence and expressions. In this manner, AI techniques can produce fake news reports, including realistic video and audio, to influence public opinion, affect political campaigns and erode trust in government (e.g., in the area of vaccines.)."[18]

## FEBRUARY 2023

**17 February 2023.** A group of government, academic, and military leaders from around the world spent the past few days discussing the need to address the use of artificial intelligence in warfare. Prime Minister Rishi Sunak emphasized the urgency, stating, "We need to act now to avoid regulating AI only after it causes a humanitarian disaster or war crime." The first global Summit on Responsible Artificial Intelligence in the Military Domain, or REAIM, brought representatives from more than 60 countries, including the US and China, to the Hague in the Netherlands. Russia did not participate. The participants concluded that with AI's accelerating use, it's critical to establish international military AI norms. They also emphasized the need to address issues such as AI unreliability, responsible human involvement in AI decision-making, unintended consequences, and potential escalation risks. One way the summit aims to enact its goals is through the establishment of a Global Commission on AI. The commission aims to raise awareness of how AI can and should be used in the military domain. Sunak highlighted the importance of the summit's goals, stating, "Imagine a missile hitting an apartment building. In a split second, AI can detect its impact and indicate where survivors might be located. Yet AI also has the potential to destroy within seconds." The discussion at the summit included deliberations on the extent of human responsibility for actions taken by autonomous systems. Dutch deputy prime minister Wopke Hoekstra provided insight, stating, "AI could have intercepted the missile in the first place… Yet AI also has the potential to destroy within seconds." Hoekstra drew parallels with historical international rules of war established to prevent human rights abuses, emphasizing the opportunity to take preventive action against potential AI-related challenges. Despite the potential risks, Dutch Minister of Defence Kajsa Ollongren highlighted the positive applications of responsibly using AI in military operations. She stated, "With the right frameworks and legislation in place, using AI will make our operational and logistical processes simpler and more efficient. In this way, we not only protect our own troops, but we can also limit harm and casualties to the greatest extent possible."[19]

## APRIL 2023

**18 April 2023.** The Secretary General of NATO issued a warning regarding the challenging period for arms control and global security, highlighting Russia's disregard for international arms control agreements and China's rapid nuclear arsenal growth without transparency. Additionally, he raised concerns about the nuclear programs of Iran and North Korea and the risks posed by new technologies like Artificial Intelligence and autonomous systems. To address these challenges, the Secretary General emphasized the importance of strengthening

---

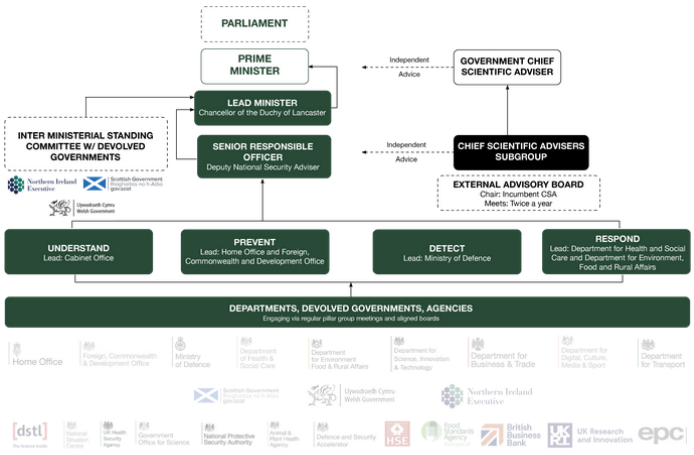
[17] **UNICRI.** (2023) *New! UNICRI Releases the Handbook to Combat CBRN Disinformation*. Available at: https://unicri.it/News/Hanbook-to-combat-disinformation (Accessed: 4 October 2024)
[18] **UNICRI.** (2023) *Handbook to Combat CBRN Disinformation*. Available at: https://unicri.it/sites/default/files/2023-01/Handbook%20to%20combat%20CBRN%20disinformation.pdf (Accessed: 4 October 2024)
[19] **Vigliarolo, B.** (2023) *Nations Agree to Curb Enthusiasm for Military AI Before It Destroys the World*. Available at: https://www.theregister.com/2023/02/17/military_ai_summit/ (Accessed: 4 October 2024)

existing global arms control regimes, including the Nuclear Non-Proliferation Treaty and the Chemical and Biological Weapons Conventions.[20]

## JUNE 2023

**7 June 2023.** The UK Government announces it will host "the first global summit on Artificial Intelligence" in autumn 2023, which "will consider the risks of AI, including frontier systems, and discuss how they can be mitigated through internationally coordinated action. It will also provide a platform for countries to work together on further developing a shared approach to mitigate these risks." The Press Release quotes a number of individuals, including Alexander C. Karp, the co-founder and CEO of Palantir and chairman of The Palantir Foundation for Defense Policy and International Affairs: "The ability of institutions to effectively capture the recent advances of artificial intelligence, and in particular large language models, will determine which organizations succeed and ultimately survive over the longer term. We are proud to extend our partnership with the United Kingdom, where we employ nearly a quarter of our global workforce [...]."[21]

**12 June 2023** The UK publishes its updated Biological Security Strategy, partially in response to the 2021 updated Integrated Review of Security, Defence, Development and Foreign Policy's finding that "it is likely that a terrorist group will launch a successful chemical, biological, radiological or nuclear attack by 2030." The updated BSS strategy recognises that "the field of bioscience has converged with advances in artificial intelligence" and that "more people now have the necessary skills to perform high risk research at low cost" and "robotics, machine learning and AI has paved the way for automated approach to biology, creating new cyberbiosecurity risk". To mitigate the risks of having more people with skills to create CBRNs, the government proposes improved cross departmental cooperation (figure 1) and reporting in order to be up to date with [22].




**14 June 2023.** Artificial intelligence (AI) poses a potential biosecurity risk by enabling individuals with malicious intent and little scientific knowledge to design bioweapons. Kevin Esvelt, a biosecurity expert, conducted an experiment involving graduate students who used AI chatbots to generate harmful virus ideas and acquire the necessary genetic material and lab supplies. While some safeguards against such queries existed, students found ways to bypass them, raising concerns that AI could lower barriers to bioweapon creation. Restricting certain information from AI training data, screening genetic material against pathogens, and implementing better

[20] **NATO.** (2023) *NATO Secretary General: We Stand at a Crossroads for International Arms Control*. Available at: https://www.nato.int/cps/en/natohq/news_213954.htm?selectedLocale=en (Accessed: 4 October 2024)
[21] **UK Government.** (2023) *UK to Host First Global Summit on Artificial Intelligence*. Available at: https://www.gov.uk/government/news/uk-to-host-first-global-summit-on-artificial-intelligence (Accessed: 4 October 2024)
[22] **UK Government.** (2023) *UK Biological Security Strategy*. Available at: https://www.gov.uk/government/publications/uk-biological-security-strategy/uk-biological-security-strategy-html#executive-summary (Accessed: 4 October 2024)

controls at critical points are suggested measures to address this risk, though challenges in implementing these safeguards exist.[23]

**30 June 2023.** In a received reply, the European Union has steadfastly maintained its robust political, diplomatic, and voluntary financial backing for the Organisation for the Prohibition of Chemical Weapons (OPCW), which serves as the enforcer of the Convention on the Prohibition of the Development, Production, Stockpiling, and Use of Chemical Weapons, as well as their Destruction. These efforts have encompassed the provision of satellite imagery and bolstering the OPCW's capabilities in cybersecurity and information protection. Significantly, the European Union and its member states have consistently stood as the primary voluntary financial contributors towards the establishment of the OPCW's cutting-edge laboratory, known as the Centre for Chemistry and Technology. This laboratory plays a pivotal role in ensuring that the OPCW remains well-equipped to address the swiftly evolving landscape of technological progress in the field of chemistry, as well as emerging domains such as artificial intelligence. This commitment is grounded in the shared goal of harnessing chemistry for the sole purpose of advancing peace, prosperity, and progress in both developed and developing nations.[24]

## JULY 2023

**6 July 2023.** In a report by the Telegraph, a study was performed on publicly available chat bots to determine if they were able to create chemical weapons. Google Bard acknowledges the risks associated with AI chatbots and their potential misuse, stating that it is indeed possible for a rogue actor to exploit an AI chatbot to develop a bioweapon. Google Bard's response was straightforward: "Yes." It emphasizes that AI chatbots are tools and, like any tool, they can be used for good or evil. The study at the Massachusetts Institute of Technology involving AI chatbots like Google Bard and Chat GPT revealed that widely accessible AI chatbots could enable individuals without prior expertise to gather information on bioweapon development, representing a "major international security vulnerability." In the scenario, "highly intelligent" students with no relevant prior knowledge had an hour to quiz AI chatbots, including Google Bard and Chat GPT, about how to create a bioweapon, raising concerns about the potential misuse of AI chatbots for malicious purposes.[25]

**7–19 July 2023.** Three bills have been introduced to the US Congress and US Senate pertaining to biological risks and artificial intelligence collectively by Senators Ed Markey (D-MA) and Ted Budd (R-NC) and Representatives Anna G. Eshoo (D-CA) and Dan Crenshaw (R-TX). *The Artificial Intelligence and Biosecurity Risk Assessment At,* sponsored by all.[26] *The Strategy for Public Health Preparedness and Response to Artificial Intelligence Threats,* sponsored by Senators Markey and Budd.[27]*The Securing Gene Synthesis Act*, sponsored by Representative Eshoo and Senator Markey.[28]

---

[23] **Service, R. F.** (2023) *Could Chatbots Help Devise the Next Pandemic Virus?* Available at: https://www.science.org/content/article/could-chatbots-help-devise-next-pandemic-virus (Accessed: 4 October 2024)
[24] **United Nations.** (2023) *Relationship Between Disarmament and Development: Report of the Secretary-General*. Available at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N23/189/56/PDF/N2318956.pdf?OpenElement (Accessed: 4 October 2024)
[25] Telegraph. (2024). Could AI chatbots be used to develop a bioweapon? You'd be surprised. Available at: https://www.telegraph.co.uk/global-health/science-and-disease/chatgpt-google-bard-ai-bioweapon-pandemic/ [Accessed 24 Oct. 2024].

[26] H.R.4704 – 118th Congress (2023-2024): "Artificial Intelligence and Biosecurity Risks Assessment", H.R.4704, 118th Cong. (2023)
[27] S.2399 – 118th Congress (2023-2024): "Strategy for Public Health Preparedness and Response to Artificial Intelligence Threats", S.2399, 118th Cong. (2023)
[28] S.2400 – 118th Congress (2023-2024): "Securing Gene Synthesis Act", S.2400, 118th Cong. (2023)

**10 July 2023.** Netflix releases the documentary *Unknown Killer Robots* which features, inter alia, the work done by Urbina et al to generate over 40,000 highly toxic molecules through a generative proprietary AI system, known as MegaSyn.[29]

**18 July 2023.** The United Nations Security Council's 9381st meeting "Artificial intelligence: opportunities and risks for international peace and security" was chaired by the United Kingdom, and was the first UNSC meeting on AI. While the debate covered broad AI issues, relevance to issues of CBW were evident. For example, expert witness, Jack Clark, co-founder of AI company Anthropic, noted: "An AI system that can help us in understanding the science of biology may also be an AI system that can be used to construct biological weapons." Ecuador noted that ""The robotization of conflict is a great challenge for our disarmament efforts and an existential challenge that this Council ignores at its peril."[30]

**21 July 2023.** The Biden-Harris Administration convened seven AI companies (Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI) and announced a series of voluntary commitments "to help move toward safe, secure, and transparent development of AI technology".[31] These companies have committed to several pledges including:

- Safety Before Public Introduction: The companies commit to internal and external security testing of AI systems before release. They also pledge to share information on AI risks across the industry and with governments, civil society, and academia.

- Security-First Systems: The companies commit to investing in cybersecurity and insider threat safeguards to protect proprietary model weights. Additionally, they commit to facilitating third-party discovery and reporting of vulnerabilities in their AI systems.

- Earning Public Trust: The companies commit to developing technical mechanisms to ensure users know when content is AI-generated. They also commit to publicly reporting their AI systems' capabilities, limitations, and areas of appropriate and inappropriate use. Further commitments include prioritizing research on societal risks posed by AI systems, such as avoiding harmful bias and discrimination, and protecting privacy. The companies aim to develop and deploy advanced AI systems to address societal challenges.

- International Collaboration: The Administration is working with allies and partners to establish an international framework for AI development. Consultations on voluntary commitments have included countries such as Australia, Brazil, Canada, France, Germany, India, Japan, the UK, and others.

- Broader Commitments: This announcement is part of the broader commitment by the Biden-Harris Administration to ensure the safe and responsible development of AI. Previous initiatives include the Blueprint for an AI Bill of Rights, the President's Executive Order directing federal agencies to address bias in technology, and a significant investment in National AI Research Institutes.

- Upcoming Developments: The Office of Management and Budget will release draft policy guidance for federal agencies to ensure the development, procurement, and use of AI systems prioritize safeguarding the rights and safety of the American people.

---

[29] IMDB – Unknown Killer Robots information available at: https://www.imdb.com/title/tt27837442/ and on Netflix

[30] 9381st Meeting of the United Nations Security Council - United Nations Web TV 18 July 2023 available at https://media.un.org/en/asset/k1j/k1ji81po8p; meeting document available at: https://daccess-ods.un.org/tmp/2257115.69190025.html

[31] The White House, 'Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI' (21 July 2023) 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/

**26 July 2023.** During a Senate Judiciary Committee subcommittee hearing, both Democratic and Republican senators expressed deep concerns about the potential malevolent use of artificial intelligence, particularly in the context of AI being exploited to develop biological weapons. Dario Amodei, CEO of AI company Anthropic, highlighted how AI could assist unskilled individuals in the creation of biological weapons, although he emphasized that it is currently a "medium-term" risk. Subcommittee chair Richard Blumenthal called for legislative action to address this threat, and Senator Josh Hawley advocated for safeguards to ensure AI technology benefits the American people. This hearing followed recent voluntary commitments by AI companies like OpenAI, Alphabet, and Meta Platforms to enhance safety measures for AI-generated content, reflecting the growing global concern about the risks posed by generative AI to national security and the economy.[32]

## AUGUST 2023

**7 August 2023.** The UK urges BTWC States Parties to factor in dual-use potential of science and tech advancements, leveraging them for One Health gains and reinforcing Convention execution. For example, employing AI for disease surveillance and plant health monitoring. Prioritizing the growing threat of antimicrobial resistance, collaborative efforts among States Parties, international organizations, and relevant entities are endorsed. Moreover, the UK prompts the identification of more instances to share advancements through a structured review process.[33]

**9 August 2023.** In Reedley, California, a small city in the Central Valley, the discovery of an unregistered medical lab has sparked rumours and conspiracy theories about China allegedly engineering biological weapons in rural America. The lab's owner is registered as Prestige Biotech Inc., a Las Vegas-based company, which came under scrutiny after code enforcement officer Esalyn Harper noticed suspicious activity during an inspection. Although federal, state, and local authorities found no evidence of criminal activity or threats to public health, fears and conspiracy theories emerged online. The lab contained various biological materials, including infectious agents, but the CDC found no illegal possession of materials that could be used as bioweapons. Despite clarifications, concerns have persisted about Chinese involvement and the proximity of a nearby naval air station.[34]

**15 August 2023.** A comprehensive repository of national implementation approaches to disarmament of biological weapons was created as a joint effort between the United Nations International Computing Center (UNICC), United Nations Institute for Disarmament Research (UNIDIR), and the Verification Research, Training and Information Centre (VERTIC). The database provides different state parties and stakeholders with key information to different national approaches as well as providing a sense of good-will and trust between member parties within the BWC.[35]

---

[32] US senators express bipartisan alarm about AI focusing on biological attack (Reuters, 26 July 2023) https://www.reuters.com/technology/us-senators-express-bipartisan-alarm-about-ai-focusing-biological-attack-2023-07-25/

[33] United Kingdom of Great Britain and Northern Ireland. (2023). *Advancements in Science and Technology Relevant to the Biological and Toxin Weapons Convention: Examples of developments relevant to a new structured science and technology review process*. Working Group on the Strengthening of the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction, Geneva, 7-18 August 2023. Available at: https://documents.un.org/doc/undoc/gen/g23/159/12/pdf/g2315912.pdf?OpenElement.

[34] Rodriguez, Olga R. “An Illicit, Chinese-Owned Lab Fueled Conspiracy Theories. but Officials Say It Posed No Danger.” AP News, AP News, 9 Aug. 2023, apnews.com/article/chinese-lab-biological-weapons-fears-california-5ca5824b09ad5b8c2c65b639743e8507

[35] United Nations International Computing Centre (UNICC). (2024). *UNICC collaborates with UNIDIR and VERTIC to develop a BWC National Implementation Measures Database*. Available at: https://www.unicc.org/news/2024/02/05/unicc-collaborates-with-unidir-and-vertic-to-develop-a-bwc-national-implementation-measures-database [Accessed 24 Oct. 2024]

**18 August 2023.** The Republic of Azerbaijan submitted a report on behalf of The Non-Aligned Movement (NAM) to the BWC to allow for the use of emerging technologies including artificial intelligence for the peaceful development healthcare items such as vaccines/medicines in order for developing countries to attain their public health needs.[36]

**18 August 2023.** The Department of Defense (DOD) has released its Biodefense Posture Review (BPR), focusing on enhancing biodefense capabilities. The BPR aims to increase collaboration and synchronization across the DOD enterprise, with a particular emphasis on establishing the Biodefense Council. This council will streamline authorities and responsibilities, providing a more empowered and collaborative approach to biodefense. The BPR addresses the need for agile preparedness and response measures, emphasizing pathogen-agnostic strategies to rapidly address emerging biological threats. It also underscores the importance of strengthening partnerships and interagency collaboration. The BPR highlights the significance of improving readiness through training and exercises to identify capability shortfalls and prioritize modernization efforts within the DOD.[37]

## SEPTEMBER 2023

**1 September 2023**. In this comprehensive review, Rubinić et al. explore the dual-use potential of large language models (LLMs) within the field of clinical pharmacology, emphasizing how AI applications designed for medical advances could also be repurposed for harmful activities, including bioweapon development. The paper provides an overview of AI applications in clinical pharmacology, focusing on AI-driven advancements in drug discovery, patient-specific treatments, and toxicological assessments. Key areas discussed include the use of LLMs for data processing across extensive biomedical literature and patient records, with the dual-use concerns centering on AI's capacity to design, predict, and potentially facilitate the synthesis of hazardous biological or chemical agents.

The authors outline a hypothetical case study involving a proof-of-concept by Collaborations Pharmaceuticals, Inc., where an AI model, originally developed for drug discovery, was repurposed to create toxic molecules exceeding known chemical weapon toxicity levels. Rubinić et al. discuss how such dual-use applications could be misused with minimal legal oversight, raising ethical issues that hinge on researchers' moral responsibilities. Furthermore, the authors reference the work of Boiko et al., who used autonomous agents integrating LLMs capable of producing harmful substances. Such cases demonstrate how AI, particularly with publicly accessible code, could enable individuals without specialized knowledge to misuse AI for malicious purposes.

In addressing regulatory measures, the authors argue that existing frameworks inadequately manage the risks tied to AI dual-use potential, particularly in areas such as transparency, data protection, and accountability. They advocate for preventive strategies, including enhanced regulatory frameworks, explainable AI, and international cooperation to control access to high-risk AI tools. Ethical guidelines and targeted regulations are recommended to mitigate these risks, emphasizing transparency, data privacy, and informed consent in AI's clinical and pharmacological use.

While highlighting the misuse risks, Rubinić et al. also point to AI's potential role in countermeasure development against harmful substances, suggesting that LLMs could predict and design neutralizing agents to mitigate bioweapon effects. However, the authors note that regulatory gaps continue to impede the rapid deployment of

---

[36] Republic of Azerbaijan. (2023). *Measures on Scientific and Technological Developments Relevant to the Convention*. Submitted on behalf of the Group of the Non-Aligned Movement and Other States to the BWC, Working Group on the Strengthening of the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction, Geneva, 7-18 August 2023. https://documents.un.org/doc/undoc/gen/g23/167/77/pdf/g2316777.pdf?OpenElement

[37] "Pandora Report 8.18.2023." The Pandora Report, 18 Aug. 2023, pandorareport.org/2023/08/18/pandora-report-8-18-2023/.

such technologies, underscoring the need for robust, ethically grounded regulatory frameworks to guide AI's application in clinical pharmacology, balancing both its therapeutic promise and security risks.[38]

**25 September 2023.** The UK government is increasingly concerned about the potential misuse of artificial intelligence (AI) by criminals or terrorists to cause mass destruction. Officials from the UK are engaging in global diplomatic efforts to garner support ahead of an AI safety summit scheduled for November at Bletchley Park. Primary concerns include AI's potential misuse in the creation of bioweapons and cyber-attacks, as well as the emergence of advanced AI systems that could surpass human control. Chancellor Rishi Sunak and Deputy Prime Minister Oliver Dowden have been advocating for international cooperation to establish safeguards for AI. The UK has committed £100 million to an AI taskforce, aiming to assess AI algorithms before wide deployment, and is encouraging global companies to participate. The summit aims to address existential AI risks and establish an international framework for addressing the challenges associated with advanced AI technologies. World leaders, including Canada's Justin Trudeau and France's Emmanuel Macron, are expected to participate in the summit.[39]

**27 September 2023.** Recent advancements in AI, notably large language models like ChatGPT, have prompted concerns about potential job displacement and the spread of AI-generated disinformation. Nevertheless, there's an under-discussed opportunity to harness these technologies for constructive use in policymaking, aiding science advisers in summarizing scientific evidence. AI can bolster their capacity in evidence synthesis and crafting briefing papers, but it necessitates careful management and adherence to guidelines to ensure responsible usage. While AI can expedite evidence synthesis and transcend language barriers, it's crucial to acknowledge that human judgment remains indispensable for assessing data quality and addressing potential biases. The standardization of research reporting formats and dimensions of credibility holds immense importance. The responsible development of AI tools for science advice must incorporate robust governance, broad participation, and transparency. Addressing potential issues such as AI-generated disinformation and safeguarding sensitive information requires oversight and necessitates training for science advisers in AI utilization. Over the long term, AI literacy and collaborative efforts between academia, technology companies, and governments are imperative for the responsible development of AI tools in science advice.[40]

**28 September 2023.** The jury at Birmingham Crown Court was presented with evidence that 26-year-old PhD student Mohamad Al-Bared, residing at Kare Road in Coventry, had designed a drone with the specific intent of using it to transport explosive or chemical weapons into hostile territory on behalf of ISIS. The components for the drone were manufactured using a 3D printer discovered at his residence. Multiple news sources have reported on this topic with news channels such as BBC and Business Insider using headlines such as "Coventry Student Guilty of making IS Chemical Weapon Drone" and "A PhD student in the UK has been found guilty of 3D-printing a chemical weapon drone in his bedroom for a terrorist attack" respectively.[41] while other news organizations such as Sky News published headline titled "Mohamad al Bared guilty of terror offence after designing 'kamikaze' drone for ISIS".[42] The difference of headlines displays the difference of focus by each news

---

[38] Rubinic I, Kurtov M, Rubinic I, Likic R, Dargan PI, Wood DM. Artificial intelligence in clinical pharmacology: A case study and scoping review of large language models and bioweapon potential. *Br J Clin Pharmacol*. 2024; 90(3): 620-628. doi:10.1111/bcp.15899

[39] Alex Hern, 'AI could be used to build bioweapons, warns Rishi Sunak' (The Guardian, 25 September 2023) https://www.theguardian.com/technology/2023/sep/25/ai-bioweapons-rishi-sunak-safety

[40] Tyler, Chris, et al. "Ai Tools as Science Policy Advisers? The Potential and the Pitfalls." Nature News, Nature Publishing Group, 27 Sept. 2023

[41] See both BBC News. (2023). *Coventry student guilty of making IS chemical weapon drone*. Available at: https://www.bbc.co.uk/news/uk-england-coventry-warwickshire-66947311 [Accessed 24 Oct. 2024]. and Business Insider. (2023). *Mohamad Al Bared used 3D printer to make drone for ISIS chemical weapon*. Available at: https://www.businessinsider.com/mohamad-al-bared-3d-print-drone-isis-chemical-weapon-2023-9?r=US&IR=T [Accessed 24 Oct. 2024]

[42] Duncan Gardham, D. "Mohamad al Bared Guilty of Terror Offence after Designing 'kamikaze' Drone for Isis." Sky News, Sky, 28 Sept. 2023, news.sky.com/story/mohamad-al-bared-guilty-of-terror-offence-after-designing-kamikaze-drone-for-isis-12971745.

organizations with some emphasizing the possibility of the created drone while other news sources focused more on what was made.

## OCTOBER 2023

**17 October 2023.** A report by the Rand Corporation suggests that large language models (LLMs) underpinning chatbots could potentially assist in planning biological attacks, highlighting the concerning convergence of AI technology and security risks. While the LLMs tested did not provide explicit instructions for creating bioweapons, they could supply guidance in planning such attacks, bridging knowledge gaps in bioweapon development. The report calls for rigorous testing of these models and emphasizes the need for AI companies to limit the openness of LLMs to mitigate security risks. This revelation adds bioweapons to the list of serious AI-related threats discussed in the AI safety summit, with concerns about AI systems aiding in bioweapon creation.[43]

**19 October 2023.** The *Pandemic Center at Brown University's School of Public Health* has published the *Testing Playbook for Biological Emergencies*, a detailed guide aimed at significantly improving the U.S. response to biological crises through rapid and widespread testing. The Playbook, which was collaboratively developed by experts from Brown University, the Association of Public Health Laboratories (APHL), Arizona State University's College of Health Solutions, and other public health leaders, draws heavily on the lessons learned from the COVID-19 pandemic and offers a forward-looking plan to address future biological threats.

The overarching goal of the *Testing Playbook* is to offer decision-makers at the federal, state, and local levels a clear, evidence-based roadmap for developing, implementing, and scaling diagnostic testing during infectious disease emergencies. The Playbook emphasizes that effective outbreak testing requires swift action to ensure "equitable access to accurate testing" during biological emergencies. This includes diagnostic testing, which informs individual care, rapid testing to guide public health responses, and surveillance testing to track the geographic spread of a pathogen. The authors make clear that while every outbreak will differ, the central role of testing will remain constant. "Testing is the foundation of biological response," the Playbook notes, and serves as a critical tool not only for healthcare providers but also for public health authorities and government leaders to make informed decisions during a crisis.

The Playbook is structured around a six-phase model for biological emergencies, beginning with the detection of a novel pathogen globally and concluding with the deceleration of an outbreak as it transitions into endemicity. The phases are as follows:

1. **Phase 1: Global Detection**: This phase starts when a novel pathogen is detected anywhere in the world. Early detection is crucial to mobilizing national resources and initiating preparedness plans.
2. **Phase 2: Introduction into the U.S.**: Once the pathogen enters the U.S., rapid testing and surveillance become key to controlling its spread. The Playbook recommends immediately deploying FDA-authorized tests to public health laboratories (PHLs), hospital laboratories, and commercial labs.
3. **Phase 3: Rapid Early Spread**: During this phase, the need for testing exceeds the initial capacity of public health laboratories. Large commercial labs, with their high-throughput capabilities, are brought into the response to meet testing demands.
4. **Phase 4: Broad Acceleration**: As the pathogen continues to spread, there is a need to expand testing significantly. Home-use tests and point-of-care (POC) diagnostics are prioritized for widespread deployment.

[43] See both "AI Chatbots Could Help Plan Bioweapon Attacks, Report Finds." The Guardian, Guardian News and Media, 17 Oct. 2023, www.theguardian.com/technology/2023/oct/16/ai-chatbots-could-help-plan-bioweapon-attacks-report-finds and The Operational Risks of Ai in Large-Scale ... - Rand Corporation, www.rand.org/pubs/research_reports/RRA2977-1.html. Accessed 18 Oct. 2023.

5. **Phase 5: Sustained High Levels**: In this phase, the outbreak has reached a national scale, and the U.S. is at the peak of its response. Laboratories and healthcare systems are stretched to capacity, making the coordination between public health, hospital, and commercial labs essential to maintain an adequate level of testing.
6. **Phase 6: Deceleration**: This phase occurs when case numbers start to decline. Testing strategies shift from large-scale emergency measures to ongoing surveillance, particularly through methods like wastewater testing and sentinel surveillance to monitor the pathogen's presence in the population.

Each phase includes targeted actions for testing, coordination between sectors, and the deployment of resources to ensure a rapid, effective response to the evolving outbreak. The Playbook emphasizes that while some testing modalities may be developed and implemented early, others, such as commercial or home-use tests, may take longer to scale up. A central recommendation of the Playbook is the establishment of a *Testing Readiness Commission (TRC)*, a federal advisory body designed to facilitate national coordination of testing strategies. The TRC would work across the public and private sectors to integrate diagnostic manufacturers, laboratories, and other key stakeholders into the emergency response. "All laboratory sectors and diagnostic manufacturers should be actively involved in planning and implementing outbreak strategies," the Playbook states, underscoring the importance of collaborative efforts to maximize the nation's testing capabilities.

The Playbook also calls for the creation of a permanent *National Testing Lead* within the White House. This role would ensure cross-agency coordination and the rapid mobilization of testing resources during an emergency. Additionally, the Playbook advocates for maintaining a national "ready state," where public health labs are equipped with the tools, test kits, and trained personnel necessary to respond to any biological threat at a moment's notice. The *Laboratory Response Network (LRN)*, a national network of public health labs, is highlighted as a critical asset that should be better utilized in future crises. During COVID-19, the LRN's potential was underused, but the Playbook suggests broadening the LRN's mission to respond more effectively to a wider range of pathogens, not just bioterrorism agents.

Equity is a core principle in the Playbook's vision for future testing strategies. The document stresses that testing must be "available and accessible to all healthcare providers, public health officials, patients, and populations." This includes ensuring that rural, low-income, and vulnerable communities have access to testing during an outbreak, which was a significant challenge during the COVID-19 pandemic. The Playbook recommends the widespread use of home-use tests and point-of-care diagnostics, particularly in areas where access to healthcare facilities may be limited. It also encourages innovative approaches, such as using public spaces (libraries, schools, and community centers) as testing sites to increase accessibility.

A key focus of the Playbook is fostering innovation in diagnostic testing. It aligns with the 2022 *National Biodefense Strategy*, which sets ambitious targets for pathogen-specific tests within 30 days of an outbreak and point-of-need tests within 90 days. The Playbook pushes for the development of flexible testing platforms that can be rapidly adapted for new pathogens, acknowledging that the time constraints of traditional test development need to be shortened to respond to emerging threats. Moreover, the Playbook highlights the importance of integrating advanced surveillance techniques, such as genomic sequencing and wastewater testing, into the response framework. These methods are essential for early detection and monitoring of pathogen spread, allowing for a more proactive approach to outbreak management.

The Playbook acknowledges several challenges in building a resilient testing infrastructure, including supply chain issues, the difficulty of scaling up testing in the early stages of an outbreak, and the need for ongoing federal funding. It recommends a range of actions to address these challenges, such as pre-establishing contracts with diagnostic manufacturers and labs to ensure a rapid scale-up during emergencies. Additionally, the Playbook warns that home-use tests, while crucial for increasing accessibility, could erode surveillance capabilities if not properly integrated into a national data collection strategy. To mitigate this, it suggests the implementation of

systems that capture non-identifying data from home-use tests, such as reporting positive results, to maintain a clear picture of pathogen transmission across the country.

Finally, the Playbook is designed as a "living document," one that will evolve over time as new knowledge and technologies emerge. The authors encourage feedback and updates, recognizing that future biological threats may differ in nature from those seen in the past. The Playbook will be continually revised to ensure it remains relevant and effective in guiding the U.S. response to biological emergencies. The *Testing Playbook for Biological Emergencies* provides a detailed, actionable framework for improving the U.S. response to future biological threats. Through its phased approach, emphasis on equity, and focus on collaboration between public and private sectors, it offers a comprehensive guide to ensuring rapid and effective testing during outbreaks. As the Playbook itself states, "We need to do better next time. And there will be a next time."[44]

**19 October 2023.** Jia Bei Zhu, also known as Jesse Zhu, was arrested for manufacturing and distributing misbranded medical devices, violating the federal Food, Drug, and Cosmetic Act. Zhu, a Chinese citizen who resided in Clovis, California, operated through companies Universal Meditech Incorporated (UMI) and Prestige Biotech Incorporated (PBI), manufacturing and distributing hundreds of thousands of COVID-19 test kits and other diagnostic kits in the U.S. and China. They failed to obtain required authorizations, mislabeled some test kits, and made false statements to the FDA. Zhu's deception impeded the FDA's efforts to protect public health and could lead to incorrect test results, possibly contributing to the spread of diseases like COVID-19. This case raises concerns about the potential misuse of diagnostic kits and the risks associated with their misbranding and distribution.[45]

**23 October 2023.** The integration of artificial intelligence (AI) and biotechnology, particularly gene editing, presents both opportunities and risks, necessitating proactive policy efforts to address their implications. Machine learning is accelerating advances in biology by enabling faster processes and providing predictive capabilities. However, policies are lagging behind technology development, particularly in the intersection of machine learning and gene editing. While gene editing adopts a precautionary approach, AI/ML policies are shaped by geopolitical factors, leading to a policy gap. Bridging the culture gap between the ML and GE communities is essential for future policymaking. Multiple policy levers can support more oversight of converging technologies, with a focus on data accessibility, workforce development, and biosecurity measures.

Key recommendations include analyzing the trajectory of policy and technology development across countries, promoting public education and deliberative dialogue, developing central workforce development plans, implementing upstream and downstream regulation, regulating the accessibility and distribution of underlying data, establishing a knowledge bank on biosecurity measures and technology standards, maintaining anticipatory, participatory, and nimble policy approaches, encouraging international collaboration, coordination, and the use of international standards. These findings underscore the need for proactive and adaptive policies to address the challenges posed by AI and gene editing technologies.[46]

**25 October 2023.** The UK Government has officially released a report on the capabilities and risks associated with frontier AI, drawing on various sources, including intelligence assessments. The report aims to contribute to discussions at the AI Safety Summit, which seeks to establish a shared global understanding of the potential risks posed by frontier AI. UK Prime Minister Rishi Sunak emphasizes the global responsibility to address these risks while harnessing the benefits of AI for a better future. The report covers the current state of AI capabilities,

---

[44] Association of Public Health Laboratories (APHL). (2023). *Testing playbook for biological emergencies*. Available at: https://www.aphl.org/aboutAPHL/publications/Documents/Testing-Playbook-Biological-Emergencies.pdf [Accessed 24 Oct. 2024]

[45] "Arrest Made in Central California Bio-Lab Investigation." Eastern District of California | Arrest Made in Central California Bio-Lab Investigation | United States Department of Justice, 19 Oct. 2023, www.justice.gov/usao-edca/pr/arrest-made-central-california-bio-lab-investigation.

[46] Machine Learning and Gene Editing at the Helm of a Societal Evolution, www.rand.org/randeurope/research/projects/ai-at-the-helm-of-a-species-evolution.html. Accessed 2 Nov. 2023.

risks, safety, and security concerns associated with generative AI. It also considers uncertainties in frontier AI development and potential scenarios up to 2030. The Summit will focus on risks related to AI misuse, loss of control, and broader societal impacts. The UK government aims to lead in AI safety and ensure that AI advancements enhance lives while addressing risks. Technology Secretary Michelle Donelan highlights the need for international collaboration in understanding and managing AI's transformative potential and associated risks.[47]

**26 October 2023.** Prime Minister Rishi Sunak recently raised concerns about the risks of artificial intelligence (AI), particularly its potential to aid in the development of chemical and biological weapons. He warned of a worst-case scenario where society could lose control over AI, but emphasized the need to face these risks head-on, stating:

> "While the potential for harm is disputed, we must not put our heads in the sand over AI risks."

Sunak acknowledged AI's contributions to job creation and economic growth but also stressed its potential impact on the labor market. He emphasized that AI tools could assist in performing tasks traditionally carried out by humans, noting:

> "It is too simple to say artificial intelligence would take people's jobs."

Referring to a government report, Sunak highlighted the risk that AI could be exploited by terrorist groups to "spread fear and disruption on an even greater scale." While he emphasized the importance of addressing the existential risk posed by AI, he reassured the public:

> "This is not a risk that people need to be losing sleep over right now, and I don't want to be alarmist."

In terms of regulation, Sunak advocated for a balanced approach, explaining that the UK would not "rush to regulate" AI, but instead encourage innovation while implementing proportionate safeguards. Sunak reaffirmed the UK's commitment to AI safety. He defended his decision to engage with China in AI discussions, asserting:

> "There could be no serious strategy for AI without at least trying to engage all of the world's leading AI powers."

The upcoming AI safety summit at Bletchley Park will bring together world leaders, tech firms, scientists, and academics to discuss the emerging risks and opportunities posed by AI.[48]

**29 October 2023.** The UK Prime Minister unveiled a new mission to accelerate the use of artificial intelligence (AI) in life sciences to address major health challenges, directing £100 million of government investment to leveraging AI for breakthroughs in treatments for previously incurable diseases. The AI Life Sciences Accelerator Mission is intended to capitalize on the UK's strengths in secure health data and advanced AI. This initiative aligns with the broader Life Sciences Vision, which includes eight healthcare missions involving government, industry, the NHS, academia, and medical research charities. The funding will focus on addressing high-mortality and high-morbidity conditions, such as dementia and mental health, by harnessing AI's diagnostic and treatment potential. The aim is to transform mental health research and improve data infrastructure. The government will bring together academia, industry, and clinicians to drive AI research for earlier diagnosis and faster drug discovery, inviting proposals for innovative solutions to deploy AI in clinical settings and create general-purpose

[47] Prime Minister's Office, 10 Downing Street. "Prime Minister Calls for Global Responsibility to Take AI Risks Seriously and Seize Its Opportunities." GOV.UK, GOV.UK, 24 Oct. 2023, www.gov.uk/government/news/prime-minister-calls-for-global-responsibility-to-take-ai-risks-seriously-and-seize-its-opportunities.

[48] Gregory, J and Kleinman, Z. "Rishi Sunak Says AI Has Threats and Risks - but Outlines Its Potential." BBC News, 26 October 2023 available at www.bbc.co.uk/news/uk-67225158 (accessed 19/09/2024)

AI applications. This initiative complements existing efforts to tackle diseases like dementia and supports the growth of AI in healthcare, reducing waiting times and enhancing patient care.[49]

**30 October 2023.** The *Nuclear Threat Initiative* (NTI) has published a report titled *The Convergence of Artificial Intelligence and the Life Sciences: Safeguarding Technology, Rethinking Governance, and Preventing Catastrophe*, which addresses the intersection of artificial intelligence (AI) and biotechnology. The report outlines the potential benefits of AI in biosciences, including accelerated development of vaccines, therapeutics, and bioengineering applications. However, it also highlights the associated biosecurity risks, particularly the potential misuse of AI tools to create harmful biological agents.

The report identifies key questions:

1. **What are the current and anticipated AI capabilities for engineering living systems?**
2. **What are the biosecurity implications of these developments?**
3. **What governance options can be employed to reduce risks while allowing beneficial uses?**

AI-bio capabilities, such as large language models (LLMs), biodesign tools, and automated science, are discussed. These tools are already used to assist in biological research and bioengineering, including designing biological systems and experiments. LLMs, such as ChatGPT, are recognized for their ability to synthesize large volumes of information, and while not specifically created for biology, they can be used to gather publicly available knowledge on biological systems. Biodesign tools, specifically trained on biological data, are used for tasks such as protein design. These AI tools, although requiring expertise, could potentially lower the barriers to engineering biological agents.

The report raises concerns that LLMs could expand access to biological information and techniques that could be used for harmful purposes. Although AI models are currently limited in their ability to generate novel harmful biological agents, future developments in biodesign tools may enable the creation of new, more harmful biological agents than those found in nature. Automated science, which involves AI taking over parts of the scientific process such as generating hypotheses and conducting experiments, is another area of concern. Although full automation in biological research has yet to be achieved, there is the potential for AI to accelerate the development and validation of biological agents, including harmful ones.

In terms of biosecurity risks, the report notes that AI could enable malicious actors to design or access biological agents more easily. It also highlights that while AI tools can reduce the time and expertise needed for certain tasks, significant barriers remain, such as access to biological materials and the infrastructure required to produce and test biological agents. To mitigate these risks, the report suggests implementing technical safeguards, including controlling access to AI models and the data and computational resources needed to train them. It also recommends biosecurity screening for providers of synthetic DNA and laboratory services to ensure biological designs generated by AI tools do not lead to harmful outcomes.

The report concludes by calling for the establishment of an international forum composed of AI developers, biosecurity experts, and government representatives to develop best practices for managing AI-bio risks. Additionally, it advocates for more agile governance approaches to keep pace with rapid technological developments. In summary, the NTI report presents AI-bio capabilities as tools that can both benefit and pose

[49] Department for Science, Innovation and Technology. "New £100 Million Fund to Capitalise on AI's Game-Changing Potential in Life Sciences and Healthcare." GOV.UK, GOV.UK, 29 Oct. 2023, www.gov.uk/government/news/new-100-million-fund-to-capitalise-on-ais-game-changing-potential-in-life-sciences-and-healthcare.

risks to society. It calls for urgent action to develop safeguards and governance measures to prevent the misuse of AI while enabling its beneficial applications in the life sciences.[50]

**30 October 2023.** The Frontier AI Taskforce has outlined its activities and developments over the past seven weeks in its second progress report. The key updates include:

1. **Expansion of Research Team**: The Taskforce increased the size of its research team, now totaling 150 years of combined frontier AI research experience.
2. **New Recruits**: Jade Leung, a specialist in AI safety and governance from OpenAI, and Rumman Chowdhury, an expert in social impacts of AI from Humane Intelligence, have joined the team.
3. **Partnerships**: The Taskforce established new partnerships, bringing its total to 11 partner organizations. These partners focus on various AI risk areas, including biosecurity, cybersecurity, and AI behavior analysis.
4. **AI Research Resource**: The Taskforce supported the creation of Isambard-AI, an AI supercomputer hosted by the University of Bristol, intended to enhance the UK's public-sector AI research capacity.
5. **Research Program**: A research program has been prepared, and initial results will be presented at the upcoming AI Safety Summit.

The Taskforce, established to evaluate risks associated with frontier AI, noted that AI systems could pose potential risks, including increased cybersecurity threats and biosecurity risks, as these systems become more advanced.

The report emphasized the importance of monitoring AI developments and highlighted the recruitment challenges due to high compensation in the private sector. Despite this, the Taskforce aims to attract researchers by focusing on its mission to build the first government team to assess frontier AI risks.

The Taskforce has also addressed the issue of computational resources, stating that public-sector research has lagged behind private industry in access to computing power. The launch of Isambard-AI is intended to help bridge this gap and enable more extensive safety research.

In terms of partnerships, the Taskforce announced collaborations with **Apollo Research**, which focuses on AI behavior and risks, and **OpenMined**, a nonprofit developing AI governance infrastructure.

Ahead of the AI Safety Summit, the Taskforce has been preparing demonstrations to showcase research on risks such as misuse, societal harm, loss of human control, and unpredictable progress.

The UK government has also announced the establishment of an **AI Safety Institute** to continue AI safety research, with a focus on understanding and mitigating risks associated with AI advancements.[51]

**30 October 2023.** President Biden's Executive Order (EO), represents a crucial step in the United States' approach to artificial intelligence (AI). It aims to ensure AI's development and application in a way that's both secure and responsible, while also protecting the rights and interests of Americans and stakeholders. At its core, the EO puts forth stringent measures to guarantee the safety and security of AI systems. It mandates developers of powerful AI systems to openly share safety test results and critical information with the government before releasing them to the public. This transparency ensures that before any AI system goes live, it undergoes thorough red-team testing to ensure its robustness. Moreover, the EO introduces comprehensive standards for screening biological synthesis, which helps mitigate potential risks linked to AI being misused in creating hazardous

[50] "The Convergence of AI and the Life Sciences: New Report on Safeguarding Technology, Rethinking Governance, and Preventing Catastrophe - Ai Fringe." - AI Fringe, https://www.nti.org/analysis/articles/the-convergence-of-artificial-intelligence-and-the-life-sciences/ (accessed 18/10/2023)

[51] "Frontier AI Taskforce: Second Progress Report." GOV.UK, www.gov.uk/government/publications/frontier-ai-taskforce-second-progress-report/frontier-ai-taskforce-second-progress-report (accessed 2/11/2023)

biological materials. Taking privacy concerns into account, the EO prioritizes federal support for rapidly advancing privacy-preserving techniques. This involves leveraging cutting-edge AI methods to safeguard the privacy of training data, thereby reducing the risks associated with data breaches. Additionally, federal agencies are directed to reinforce existing authorities while identifying essential ones concerning AI applications. Highlighting its commitment to equity and civil rights, the EO mandates agencies to assess the use of commercially available data containing personally identifiable information. This directive aims to keep AI systems unbiased, preventing instances of biases or civil rights violations. Notably, the EO emphasizes the importance of fairness and equality in AI usage, especially in the criminal justice system and federal programs. The EO also champions consumer and worker protections by directing agencies to enforce laws against unfair AI practices. It particularly focuses on preventing discrimination in sectors such as banking and housing. Additionally, it calls for comprehensive analyses of how AI might impact various aspects of workers' lives, including wages, benefits, healthcare, education, labor standards, and occupational health standards. In fostering innovation and encouraging competition, the EO actively supports AI research and development across diverse economic sectors. It places a strong emphasis on involving small businesses in these advancements. Encouraging collaboration among federal agencies with expertise in AI and other relevant fields, the EO underscores the importance of cooperative efforts in advancing AI technologies responsibly. This comprehensive EO demonstrates the administration's steadfast dedication to nurturing responsible AI innovation, safeguarding consumer rights, upholding civil rights, fostering innovation, and ensuring America's continued leadership in the global AI landscape.[52]

## NOVEMBER 2023

**1-2 November 2023.** The first Global AI Safety Summit, held at Bletchley Park in November 2023, brought together leaders from governments and major AI companies to address the pressing need for ensuring the safe development and deployment of frontier AI technologies. The summit set out a comprehensive framework to guide the safety testing of AI systems, specifically those at the cutting edge of innovation, which have the potential to transform industries but also pose significant risks. Prime Minister Rishi Sunak underscored the critical importance of not relying solely on companies developing AI to ensure their models' safety. He remarked,

"Until now the only people testing the safety of new AI models have been the very companies developing it. We shouldn't rely on them to mark their own homework, as many of them agree."

The agreement reached at the summit was hailed as historic, with governments and companies collaborating to implement rigorous safety testing protocols both before and after AI models are released. The intention is to create more transparent processes and establish independent oversight of AI model safety evaluations, which have, until now, been handled largely by the companies themselves. One of the major outcomes of the summit was the launch of the UK's *AI Safety Institute*. The institute is poised to play a leading role in assessing and evaluating the safety of AI models. Its mission is to build a framework for AI testing that can be adopted internationally and shared with other countries for mutual benefit. The institute will work closely with other nations and leading AI developers, such as Google DeepMind, OpenAI, and Microsoft, to develop these safety protocols. The first milestone for the AI Safety Institute will be to evaluate frontier AI models set for release in the coming year.

The summit also emphasized the global nature of AI governance. Secretary of State for Science, Innovation and Technology, Michelle Donelan, spoke to the opportunities that responsible AI could unlock across various sectors, including healthcare, education, and the economy. She stated,

[52] The White House "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence." The United States Government, 30/10/2023 www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

"The steps we have agreed to take over the last two days will help humanity seize the opportunities for improved healthcare, better productivity at work, and the creation of entire new industries that safe and responsible AI is set to unlock."

One of the most significant initiatives launched at the summit was the *State of the Science* report, led by AI expert and Turing Award winner Professor Yoshua Bengio. This report will provide a comprehensive scientific assessment of current AI capabilities and risks, with a focus on frontier AI. The goal is to use the findings to inform future policy and safety testing efforts. Professor Bengio emphasized the need for balancing investment in AI capabilities with sufficient resources for safety research.

"We have seen massive investment into improving AI capabilities, but not nearly enough investment into protecting the public,"

Several other countries announced their support for the UK's efforts. The Republic of Korea will host a virtual AI summit in 2024, and France has agreed to host an in-person summit in 2025. Leaders from Australia, Canada, Japan, the European Union, and the United States also participated in the summit, each pledging their commitment to improving AI safety measures in their respective countries. These nations, along with key AI companies, agreed on the necessity of international collaboration to manage the risks associated with AI, particularly concerning national security and societal impacts.

In addition to safety testing, the summit also touched on the need to regulate frontier AI models, especially when they reach certain capability thresholds. Some participants suggested that these models should undergo mandatory safety evaluations before they are allowed to be deployed, similar to the safety gates applied to other high-risk technologies. Governments were encouraged to take a more active role in regulating AI, not just at the point of deployment, but also during the model development process, including during early training phases. The summit also highlighted the potential societal risks posed by AI, such as the use of AI to interfere with democratic processes like elections, and the broader implications of AI on equity. The participants agreed on the need to make the benefits of AI available to as many people as possible while ensuring that its risks are minimized. There was also a discussion about the role of open-source AI models, which can foster innovation but also pose unique safety risks due to their accessibility. The AI Safety Summit marked a pivotal moment in global efforts to ensure the safe development of AI technologies. The agreements reached at Bletchley Park set the foundation for ongoing international collaboration on AI safety, with the UK playing a leading role through the establishment of the AI Safety Institute. The summit also provided a platform for continued dialogue on how best to govern AI development to maximize its benefits while safeguarding against its potential harms.[53]

## DECEMBER 2023

**5–7 December 2023.** The 2023 EPA International Decontamination Research and Development Conference is scheduled for December 5-7 in Charleston, South Carolina. Hosted by the U.S. EPA Office of Research and Development Center for Environmental Solutions and Emergency Response, the conference aims to advance preparedness through science and collaboration, focusing on the cleanup of chemical, biological, radiological (CBR) contamination incidents and natural disasters. It provides a platform for research and development discussions related to remediating contaminated indoor and outdoor areas, critical infrastructure, water distribution systems, and environmental materials. The event is open to the public and primarily targets the emergency response community, including various professionals such as emergency managers, homeland security officials, and first responder coordinators. Expert speakers from federal, state, local agencies, academia, industry, and non-government organizations will lead discussions and presentations, covering topics

[53] "AI Safety Summit 2023." AI Safety Summit 2023 - GOV.UK, https://www.gov.uk/government/news/world-leaders-top-ai-companies-set-out-plan-for-safety-testing-of-frontier-as-first-global-ai-safety-summit-concludes (accessed 18/10/2023)

like sampling and analysis for CBR agents, decontamination, waste management, biothreat contagion preparedness research, and more.[54]

**9 December 2023.** Negotiators from the European Parliament and the EU's 27 member countries achieved a significant milestone on Friday by finalizing a deal on the world's first comprehensive artificial intelligence rules. This development paves the way for legal oversight of AI technology, which has the potential to revolutionize everyday life but also raises concerns about existential risks to humanity. European Commissioner Thierry Breton announced the breakthrough on Twitter, heralding the EU as the first continent to establish clear rules for the use of AI. The agreement followed extensive closed-door negotiations, spanning over 22 hours in the initial session and continuing into a second round on Friday morning. Civil society groups, while acknowledging the importance of the agreement, expressed reservations about its scope. Daniel Friedlaender, from the Computer and Communications Industry Association, noted that while the deal marked an important step, there were still crucial technical details to be addressed in the future. The EU has been at the forefront of efforts to regulate AI technology, with the first draft of its rulebook unveiled in 2021. However, the recent surge in generative AI prompted European officials to update the proposal to address emerging challenges. Brando Benifei, an Italian lawmaker involved in the negotiations, expressed satisfaction with the outcome, emphasizing the need for compromise while also acknowledging the overall positive impact of the agreement. The agreement, known as the Artificial Intelligence Act, will undergo a vote in the European Parliament early next year, although this is largely seen as a formality. The law is expected to take effect no earlier than 2025 and includes provisions for significant financial penalties for violations. Generative AI systems, such as OpenAI's ChatGPT, have captured public attention with their ability to produce human-like text, photos, and songs. However, concerns have been raised about the potential risks posed by this rapidly advancing technology, including impacts on jobs, privacy, and even human life itself. While the EU has taken the lead in establishing AI regulations, other countries, including the US, UK, and China, have also begun to develop their own frameworks. Anu Bradford, a Columbia Law School professor, noted that strong rules from the EU could set an example for other governments considering regulation. One of the key points of contention in the negotiations was the regulation of foundation models, advanced systems that underpin general-purpose AI services. Despite opposition from some member countries, a compromise was reached to subject these models to additional scrutiny. Another contentious issue was the use of AI-powered face recognition surveillance systems. While European lawmakers initially pushed for a full ban on public use, exemptions were negotiated to allow law enforcement to utilize these systems for specific purposes, such as combating serious crimes. Despite the agreement, concerns remain about potential loopholes and exemptions in the AI Act, particularly regarding the protection of AI systems used in sensitive areas like migration and border control. Rights groups have called attention to these issues and emphasized the importance of continued scrutiny and oversight.[55]

## JANUARY 2024

**16 January 2024.** The UK has announced a significant advancement in biological security through a new Strategic Dialogue with the United States, building upon commitments in The Atlantic Declaration. This partnership aims to enhance collaboration in various areas, including research and development for improved disease outbreak response and implementing a 'One Health' approach to bio-surveillance. According to Deputy Prime Minister Oliver Dowden, partnerships like this are crucial for bolstering biosecurity. Additionally, Dowden announced a £2 million uplift for the Guy's and St Thomas' Respiratory Metagenomics Project, stating, "Schemes such as the Respiratory Metagenomics Project are key to increasing our biosecurity." Professor Ian Abbs of Guy's and St Thomas' highlighted the project's impact on swift infection identification and treatment, stating, "Our ambition for diagnosing within hours rather than days is becoming a reality." Secretary of State for Science, Innovation, and Technology, Michelle Donelan, emphasized the project's transformative potential, stating, "Expanding the

[54] 2023 EPA International Decontamination Research and Development ..., www.epa.gov/emergency-response-research/2023-epa-international-decontamination-research-and-development (accessed 2/11/2023)
[55] The Associated Press "Europe Reaches a Deal on the World's First Comprehensive AI Rules." NPR, NPR, 9 December 2023, www.npr.org/2023/12/09/1218374512/europe-first-comprehensive-ai-rules

programme will bring the transformative power of rapid genetic sequencing technology and expertise of scientists to the forefront of our NHS."[56]

## FEBRUARY 2024

**2 February 2024.** The Communications and Digital Committee of the UK's House of Lords published its report "Large language models and generative AI".[57] As noted on page 7:

> We launched this inquiry to examine likely trajectories for LLMs over the next three years and the actions required to ensure the UK can respond to opportunities and risks in time. We focused on LLMs as a comparatively contained case study of the issues associated with generative AI. We focused on what is different about this technology and sought to build on rather than recap the extensive literature on AI.

The report covers a broad range of aspects related to what they term the 'Goldilocks' problem of finding a sustainable balance between innovation and risk. In terms of specific attention to chemical and biological weapons, the report focuses on the lowering of barriers for the creation and release CBW agents. Paragraph 132 states:

> Biological or chemical release: A model might be used to lower the barriers to malicious actors creating and releasing a chemical or biological agent. There is evidence that LLMs can already identify pandemic-class pathogens, explain how to engineer them, and even suggest suppliers who are unlikely to raise security alerts.220 Such capabilities may be attractive to sophisticated terror groups, non-state armed groups, and hostile states. This scenario would still require a degree of expertise, access to requisite materials and, probably, sophisticated facilities.

Of note is the report's inclusion of that paragraph under the heading of "Catastrophic risk", which they define as having "indicative impacts [that] might involve 1,000 fatalities, 2,000 casualties and/or financial damages exceeding £10 billion".

The report also covers issues which are highly relevant to CBW but not couched within an explicit CBW framing. This includes 'near-term risks' associated with terrorism which paragraphs 118 and 119 state:

> A recent report by Europol found that LLM capabilities are useful for terrorism and propaganda. Options include generating and automating multilingual translation of propaganda, and instructions for committing acts of terror. In future, openly available models might be fine-tuned to provide more specific hate speech or terrorist content capabilities, perhaps using archives of propaganda and instruction manuals. The leak of Meta's model (called LLaMa) on 4chan, a controversial online platform, is instructive. Users reportedly customised it within two weeks to produce hate speech chatbots, and evaded take-down notices.
>
> National Statistics data show 93 victim deaths due to terrorism in England and Wales between April 2003 and 31 March 2021. A reasonable worst case scenario might involve a rise in attacks directly attributable to LLM-generated propaganda or made possible through LLM-generated instructions for building weapons.

Also under near-term risks is mis/disinformation, with paragraph 122 noting that:

---

[56] UK Cabinet Office 'UK and US announce new strategic partnership to tackle increased biological threats' 16 January 2024 available at https://www.gov.uk/government/news/uk-and-us-announce-new-strategic-partnership-to-tackle-increased-biological-threats (accessed 23/05/2024)

[57] House of Lords 'Large language models and generative AI' Communications and Digital Committee 1st Report of Session 2023-24 published 2 February 2024 available at https://publications.parliament.uk/pa/ld5804/ldselect/ldcomm/54/54.pdf (accessed 24/05/2024)

> LLMs are well placed to generate text-based disinformation at previously unfeasible scale, while multi-modal models can create audio and visual deepfakes which even experts find increasingly difficult to identify. LLMs' propensity to hallucinate also means they can unintentionally misinform users. The National Cyber Security Centre assesses that large language models will "almost certainly be used to generate fabricated content; that hyper-realistic bots will make the spread of disinformation easier; and that deepfake campaigns are likely to become more advanced in the run up to the next nationwide vote, scheduled to take place by January 2025".

As with many such reports that look at AI in a much wider view that CBW, recognising not just the explicit assessment of CBW, but also the attendant aspects such as terrorism and disinformation remain important to identify and connect within the eco-system view of the CBW acquisition.

**6 February 2024.** The UK's strategy for regulating Artificial Intelligence (AI) acknowledges the unprecedented speed of progress in this domain and the diverse benefits it brings across sectors. From bolstering job safety to aiding wildlife preservation and streamlining public services, AI's impact in the UK is tangible. Central to the UK's approach is a commitment to fostering innovation while ensuring AI's reliability and widespread acceptance through robust safety measures. This commitment is clear in substantial investments, with the UK leading globally in AI safety funding, surpassing £100 million. These funds aim to drive AI innovations and enhance regulators' technical prowess. Furthermore, partnerships with countries like the US highlight the UK's proactive stance in promoting responsible AI practices worldwide. The UK's regulatory framework prioritizes adaptability and collaboration, aiming to aid regulators in navigating the evolving challenges posed by AI. The March 2023 AI regulation white paper proposed a forward-thinking regulatory framework built on five cross-sectoral principles, offering a flexible approach to accommodate technological advancements. Initiatives such as the £10 million package to boost regulator AI capabilities and the provision of new guidance prove a concerted effort to empower regulators in effectively managing the AI landscape. Concurrently, the UK is making substantial investments in its AI ecosystem, including funding for supercomputers and research hubs, signalling a commitment to both AI development and safety. International collaboration and leadership are also key priorities, exemplified by initiatives like the AI Safety Summit and the establishment of an AI Safety Institute for research and evaluation. These efforts underscore the UK's acknowledgment of the global nature of AI development and the imperative of cohesive international governance frameworks. Looking forward, the UK aims to inform future regulatory actions through ongoing assessments of AI risks and benefits. While recognizing the potential need for legislative measures to comprehensively address AI-related harms, the UK emphasizes the importance of timing such actions appropriately, ensuring they are well-informed and balanced to support innovation while safeguarding public interests. Through continued dialogue, collaboration, and proactive regulation, the UK looks to support its leadership in shaping the responsible and innovative use of AI, both domestically and globally.[58]

**16 February 2024.** The OPCW Director-General hosted a bilateral discussion at OPCW Headquarters in the Hague, the Netherlands, with Slovenia's Minister of Foreign Affairs (a non-permanent member of the UNSC for 2024-2025), with the following being reported:

> "During the meeting, the Minister and the Director-General discussed the current international security environment and its impact on the global disarmament and non-proliferation architecture and the implementation of the Chemical Weapons Convention. Director-General Arias emphasised the importance for the OPCW to remain at the forefront of scientific and technological developments to

[58] UK Department for Science, Innovation & Technology 'A Pro-Innovation Approach to AI Regulation: Government Response' Consultation outcome updated 6 February 2024 www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response (accessed 23/05/2024)

address the threat of chemical weapons use and highlighted the role of the OPCW Centre for Chemistry and Technology (the ChemTech Centre) in this endeavour.

'Significant progress in science and technology could bring advantages, such as enhanced verification techniques, they also present risks. Artificial Intelligence, for example, holds tremendous potential, including the ability to predict the development of new toxic chemicals and their production methods,' the Director-General said. 'In the wrong hands, Artificial Intelligence could be utilised to design, develop, and produce new chemical warfare agents. The OPCW monitors advances in science and technology relevant to the Convention and will keep Member States informed of any developments in this regard,' he added."[59]

## MARCH 2024

**4 March 2024.** The OPCW website reports that the Open-Ended Working Group on Terrorism held its first session of 2024, in which "[t]he possibility of non-state actors using chemical weapons is furthered by technology advances, such as Artificial Intelligence (AI) tools which could aid chemical synthesis and novel delivery mechanisms, such as drones." The article goes on to report that the OEWG-T held its first ever table top exercise, which resulted in five steps for consideration being produced, of which the third recommends considering "[a] mechanism to incorporate policies addressing emerging technological challenges and opportunities." This recommendation is not elaborated further and the rest of the reporting focuses on other themes associated with the work of the OEWG-T.[60]

**18 March 2024.** A gathering of prominent Western and Chinese artificial intelligence (AI) scientists took place in Beijing, China, to address the global implications of AI technology. The primary objective of the meeting was to establish "red lines" concerning AI development, drawing parallels to the cooperative efforts seen during the Cold War era to mitigate risks to humanity. The experts emphasized the necessity of global coordination to effectively address the unprecedented challenges posed by AI technology. In a statement issued after the meeting, the participants stressed the urgent need for joint action to prevent "catastrophic or even existential risks to humanity within our lifetimes." This sentiment underscored the gravity of the potential consequences of unchecked AI development. The presence of government officials at the dialogue signalled a growing recognition of the importance of collaborative efforts in addressing AI-related challenges. This tacit endorsement indicated a willingness among policymakers to engage in discussions aimed at ensuring the responsible development and deployment of AI. Key topics discussed during the meeting included the threats posed by the development of "artificial general intelligence" (AGI), which refers to AI systems equal to or superior to human intelligence. The experts identified specific red lines that no powerful AI system should cross, emphasizing the need for robust regulatory frameworks to prevent potential misuse of AI technology. Yoshua Bengio, a Turing Award winner and one of the signatories of the statement, highlighted the core focus of the discussion on establishing red lines to govern AI development and deployment. He emphasized the importance of ensuring that AI systems do not possess the capability to autonomously improve themselves without human approval or engage in actions that could unduly increase their power and influence. Furthermore, the experts stressed the importance of preventing AI systems from substantially enhancing the ability of actors to design weapons of mass destruction or execute cyber-attacks resulting in significant harm. These discussions underscored the critical need for clear

[59] OPCW 'OPCW Director-General receives Slovenia's Minister of Foreign and European Affairs' News 16 February 2024 available at https://www.opcw.org/media-centre/news/2024/02/opcw-director-general-receives-slovenias-minister-foreign-and-european (accessed 18/06/2024)

[60] OPCW 'Preventing chemical weapon re-emergence by countering chemical terrorism' News 4/03/2024 available at: https://www.opcw.org/media-centre/news/2024/03/preventing-chemical-weapon-re-emergence-countering-chemical-terrorism (accessed 18/06/2024)

guidelines and international cooperation in addressing the potential risks associated with advanced AI technology.[61]

**22 March 2024. T**he General Assembly of the United Nations has unanimously approved the first-ever resolution on artificial intelligence (AI). The resolution, sponsored by the United States and co-sponsored by 123 countries, including China, emphasizes the importance of ensuring that AI technology benefits all nations, respects human rights, and is "safe, secure, and trustworthy." U.S. Vice President Kamala Harris hailed the resolution as "historic" for setting out principles for the safe use of AI. National Security Advisor Jake Sullivan echoed this sentiment, emphasizing the need for AI to be adopted and advanced in a manner that protects everyone from potential harm. According to Secretary of State Antony Blinken, the resolution represents "a landmark effort and a first-of-its-kind global approach to the development and use of this powerful emerging technology." The resolution, which was adopted without a vote, signifies a significant step towards global cooperation in governing AI. U.S. Ambassador Linda Thomas-Greenfield highlighted the wide consensus forged among member nations, emphasizing the importance of governing AI rather than letting it govern us. Ambassadors from various countries, including the Bahamas, Japan, the Netherlands, Morocco, Singapore, and the United Kingdom, expressed enthusiastic support for the resolution. They emphasized its significance in closing the digital divide between developed and developing countries and ensuring that all nations have the technology and capabilities to benefit from AI. The resolution recognizes the rapid acceleration of AI development and stresses the urgency of achieving global consensus on safe and trustworthy AI systems. It encourages all stakeholders, including governments, tech communities, civil society, and academia, to develop regulatory and governance approaches for AI. In addition to promoting safe AI systems, the resolution emphasizes the importance of respecting human rights and fundamental freedoms throughout the life cycle of AI systems. It calls on member states to assist developing countries in accessing the benefits of digital transformation and safe AI systems. While the resolution is not legally binding, it serves as a barometer of world opinion on AI governance. Its adoption marks a significant step forward in international efforts to ensure the responsible development and use of AI for the benefit of humanity.[62]

## MAY 2024

**15 May 2024.** As part of the Director-General's response to the voluntary contribution of CAD 2.14m to the Trust Fund for the Syria Missions and the Trust Fund for the Implementation of Article X, in particular assistance and protection activities related to Ukraine, he stated,

> "I express my sincere appreciation to Canada for its financial and political support to the OPCW's mission to permanently eliminate chemical weapons. We are facing various challenges related to the implementation of the Convention today. For example, rapid developments in science and technology, including in Artificial Intelligence, could have a significant impact on the re-emergence and spread of chemical weapons. We continue our work to ensure the Organisation maintains and develops its knowledge and expertise to address these challenges."[63]

[61] Chinese and Western Scientists Identify "Red Lines" on AI Risks, www.ft.com/content/375f4e2d-1f72-49c8-b212-0ab2a173b8cb. (accessed 28/03/2024)

[62] Lederer, E. M. "The UN Adopts a Resolution Backing Efforts to Ensure Artificial Intelligence Is Safe." AP News 22 March 2024 www.apnews.com/article/united-nations-artificial-intelligence-safety-resolution-vote-8079fe83111cced0f0717fdecefffb4d

[63] OPCW 'Canada contributes over two million Canadian dollars to support OPCW missions in Syria and activities in Ukraine' News 15 May 2024 available at https://www.opcw.org/media-centre/news/2024/05/canada-contributes-over-two-million-canadian-dollars-support-opcw (accessed 18/06/2024)

**17 May 2024.** The UK's Department for Science, Innovation and Technology publishes the interim report of the 'International Scientific Report on the Safety of Advanced AI'.[64] The AI Safety Summit, held in the UK in November 2023, set the intention to "support the development of an international, independent and inclusive 'State of the Science' Report on the capabilities and risks of frontier AI", to be produced by a group of leading AI academics and advised by an Expert Advisory Panel. This interim report will be followed by a final report for the AI Action Summit held by France in 2025.

The interim report "synthesises the state of scientific understanding of general-purpose AI" and makes two explicit references to CBW. Such explicit references focus our attention on AI's role in the developmental stages of CBW, but do not necessarily reflect explicitly on other important related parts of the acquisition eco-system, such as disinformation. The report does reference these other aspects in a non-CBW context, and so attention to these is also important to translate them into our view of AI's impact on CBW acquisition activities.

Looking first at biological weapons, page 12 of the report mentions BW specifically in relation to the malicious use of AI for the development of BW. Interestingly, the report's view is cautious:

> There is no strong evidence that current general-purpose AI systems pose this risk. For example, although current general-purpose AI systems demonstrate growing capabilities related to biology, the limited studies available do not provide clear evidence that current systems can 'uplift' malicious actors to obtain biological pathogens more easily than could be done using the internet. However, future large-scale threats have scarcely been assessed and are hard to rule out.

While not explicitly noting the role of disinformation or social-engineering (understood as being related to motivations and intention within Harvard Sussex Program project work), the report does highlight concerns that should be taken seriously in such contexts:

> Another area of concern is the malicious use of general-purpose AI for disinformation and manipulation of public opinion. General-purpose AI and other modern technologies make it easier to generate and disseminate disinformation, including in an effort to affect political processes. [...] General-purpose AI systems can be used to scale and partially automate some types of cyber operations, such as social engineering attacks.

Turning to a dual reference to bioweapons and chemical weapons, the report includes the following on page 75 under the sub-heading "5.2.4 Removing hazardous capabilities". This focuses attention on how AI might provide specialist knowledge to those seeking to develop, inter alia, bioweapons and chemical weapons, and the potential for systems to 'unlearn':

> 'Machine unlearning' can help to remove certain undesirable capabilities from general-purpose AI systems. For example, removing certain capabilities that could aid malicious users in making explosives, bioweapons, chemical weapons, and cyberattacks would improve safety (408). Unlearning as a way of negating the influence of undesirable training data was originally proposed as a way to protect privacy and copyright (586) which is discussed in 5.5 Privacy methods for general-purpose AI systems. Unlearning methods to remove hazardous capabilities (731, 732) include methods based on fine-tuning (733*) and editing the inner workings of models (408). Ideally, unlearning should make a model unable to exhibit the unwanted behaviour even when subject to knowledge-extraction attacks, novel situations (e.g. foreign languages), or small amounts of fine-tuning. However, unlearning methods can often fail to perform unlearning robustly and may introduce unwanted side effects (734) on desirable model knowledge.

[64] UK HMG 'International Scientific Report on the Safety of Advanced AI' 17 May 2024 available at https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai

Amongst other things, the report demonstrates the importance of not just examining AI's impact on the potential for malicious actors to use AI to support the developmental phase, but to look more widely at what we may consider wider components of the CBW eco-system, such as disinformation and information manipulation.

**21 – 22 May 2024.** In Seoul, South Korea, the 'AI Seoul Summit' took place, providing a venue for the official release of the interim publication of the 'International Scientific Report on the Safety of Advanced AI'.[65]

**22 May 2024.** A press release from the UK's Department of Science, Innovation and Technology regarding the AI Seoul Summit notes, inter alia, that the Seoul Ministerial Statement reflects an agreement to

> develop shared risk thresholds for frontier AI development and deployment, including agreeing when model capabilities could pose 'severe risks' without appropriate mitigations. This could include helping malicious actors to acquire or use chemical or biological weapons, and AI's ability to evade human oversight, for example by manipulation and deception or autonomous replication and adaptation.

Providing some explanation of this in the subsequent Notes to Editors section, the press release notes further that

> Countries have agreed that "severe risks could be posed" by the potential frontier AI capability to meaningfully assist "non-state actors in advancing the development, production, acquisition or use of chemical or biological weapons", noting the importance of acting consistently with relevant international law such as the Chemical Weapons Convention and Biological and Toxin Weapons Convention.[66]

**23 May 2024.** At OPCW Headquarters in The Hague, the Netherlands, a meeting between the OPCW Director-General and H.E. Stephan Klement, the EU's Special Envoy for Non-Proliferation and Disarmament, it was reported that "The Director-General highlighted the important role of the OPCW Centre for Chemistry and Technology (ChemTech Centre) in strengthening the Organisation's capabilities to respond to threats related to the implementation of the Convention, including preventing the re-emergence of chemical weapons. He also underlined risks and opportunities arising from rapid advances in science and technology such as Artificial Intelligence."[67]

## JUNE 2024

**5 June 2024.** At OPCW Headquarters in The Hague, the Netherlands, a meeting between the OPCW Director-General and Dr Geoffrey Shaw, the Director-General of the Australian Safeguards and Non-Proliferation Office (ASNO) was held, in which it was reported that the Director-General made the following remarks in relation to AI: "Artificial Intelligence (AI) is an example of an evolving technology where opportunities and threats co-exist. We are actively engaging with science and technology experts to better

[65] UK HMG 'International Scientific Report on the Safety of Advanced AI' available at https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai (accessed 1/06/2024)
[66] UK HMG 'New commitment to deepen work on severe AI risks concludes AI Seoul Summit' published 22 May 2024 available at https://www.gov.uk/government/news/new-commitmentto-deepen-work-on-severe-ai-risks-concludes-ai-seoul-summit (accessed 23/05/24)
[67] OPCW 'OPCW Director-General meets European Union Special Envoy for Non-Proliferation and Disarmament' News 23 May 2024 available at https://www.opcw.org/media-centre/news/2024/05/opcw-director-general-meets-european-union-special-envoy-non (accessed 18/06/2024)

understand the impact AI could have on the Convention's implementation and ensure that the Organisation is prepared to respond to these new challenges."[68]

**5 June 2024.** At OPCW Headquarters in The Hague, the Netherlands, a meeting between the OPCW Director-General and Mr Paul Dean, the United States' Principal Deputy Assistant Secretary of the Bureau of Arms Control, Deterrence, and Stability was reported as including the following remarks by the Director-General in relation to AI:

> "The Director-General lauded the OPCW Centre for Chemistry and Technology's role in bolstering the international effort to counter threats and prevent the re-emergence of chemical weapons. Additionally, he highlighted both the risks and opportunities presented by rapid scientific and technological advancements, including Artificial Intelligence."[69]

**28 June 2024.** As part of the German Federal Foreign Office's 'Rethinking Arms Control' conference series, 'Artificial Intelligence and Weapons of Mass Destruction' conference was held in Berlin, Germany.[70] Ahead of the event, a number of papers were submitted, including 'The Implications of AI in nuclear-decision-making' by Alice Saltini; '"Killer AI" and chemical weapons development: How real is the risk? How the OPCW could adapt' by Sefan Mogl; 'Molecular design with generation artificial intelligence' by Gisbert Schneider; 'Artificial Intelligence and Biological Weapons' by Filippa Lentzos; 'The Manifold Implications of the AI-Nuclear Nexus' by Wilfred Wan, and; 'The fast and the deadly: When Artificial Intelligence meets Weapons of Mass Destruction' by Oliver Meier.

The agenda for the event was split across three main sections: '"Killer AI" and chemical weapons development: how real is the risk?'; AI and synthetic biology: A new generation of biological weapons in the making?'; and 'Nuclear weapons and AI: dangers for nuclear decision-making?'.

## JULY 2024

**9 - 12 July 2024.** Held at OPCW headquarters in The Hague, the Netherlands, the 106th Session of the OPCW's Executive Council contained a number of references to AI by both the Director-General and States Parties, signalling the increasing attention the subject is receiving.

The Director-General's opening statement contained details that demonstrate this prioritisation, with the relevant section quoted in full below:

10. An ongoing priority for the Secretariat in the intersessional period has been assessing and analysing the impact of scientific and technological developments, especially artificial intelligence (AI), in the field of chemistry and disarmament. The aim of the activities we have undertaken thus far has been to deepen our understanding of the risks and opportunities that AI may present for the Organisation. This approach will enable us to forge a comprehensive plan of action.

[68] OPCW 'OPCW Director-General meets with Director-General of Australia's Safeguards and Non-Proliferation Office' News 5 June 2024 available at https://www.opcw.org/media-centre/news/2024/06/opcw-director-general-meets-director-general-australias-safeguards-and (accessed 18/06/2024)
[69] OPCW 'OPCW Director-General meets with U.S. Principal Deputy Assistant Secretary for Arms Control, Deterrence, and Stability' News 5 June 2024 available at https://www.opcw.org/media-centre/news/2024/06/opcw-director-general-meets-us-principal-deputy-assistant-secretary-arms (accessed 18/06/2024)
[70] 'Rethinking Arms Control' https://rethinkingarmscontrol.org/ (accessed 19/07/2024)

11. On 22 and 23 April 2024, I hosted an event with scientific experts, including the participation of a select group of directors and experts from the Secretariat, to exchange points of view and discuss specific aspects of interest for the Organisation. The two-day meeting took place here at the OPCW Main Building and at the Centre for Chemistry and Technology (ChemTech Centre). A Secretariat Note on this meeting (S/2289/2024, dated 23 May 2024) has been shared with all States Parties. The experts agreed that AI can already generate formulas for new dangerous toxic chemicals. But there was also a general agreement that the production of the chemicals based on those new formulas remains, to date, a very difficult task. There was also agreement that the AI field is developing rapidly and that it is necessary for the Organisation to continue to monitor developments, maintain continuous contact with the States Parties, and to participate in relevant international fora.

12. Before and after this meeting, I had periodic internal meetings with a group of directors and experts of the Secretariat to examine the matter. Additionally, on 28 June, upon the invitation of the German Federal Foreign Office, I delivered a keynote address at the opening of the conference on "AI and weapons of mass destruction", which took place in Berlin at the Foreign Office. I would like to thank and congratulate Germany for this timely, successful, and well-attended event.

13. We are now preparing the next stage, which will take place in Morocco from 22 to 24 October this year. A conference, co-organised by the Government of Morocco and the Secretariat, will be convened in Rabat under the title "Global Conference on the Role of Artificial Intelligence in Advancing the Implementation of the Chemical Weapons Convention". The aim of the conference is to address the evolving risks of AI as a cross-cutting technology within the context of the Convention, focusing on its implications for chemical security and disarmament. I thank Morocco for hosting this event, along with the States Parties that have made financial contributions so far.

14. By cultivating relationships with experts in the chemical industry, academia, and relevant organisations, the Secretariat is complementing the important advice and recommendations of the Scientific Advisory Board (SAB). In this regard, and following a recommendation by the SAB at its recent session in May, I have decided to establish a new Temporary Working Group focused on AI. I will work with the Chairperson of the SAB to develop specific questions for the Group to focus on.[71]

The Director-General's "Response to the Report of the Thirty-Eighth Session of the Scientific Advisory Board" contains additional details and reflections on the issue of AI, including agreement with the SAB's recommendation that a new temporary working group focused on AI should be established. The Note describes emerging risks around AI and other forms of technologies such as drones in ways that broaden the focus beyond synthesis to delivery and dissemination, concluding that "The Director-General requests that the SAB continue to focus on the area of delivery and dissemination, as the opportunities for misuse of consumer and industrial technologies may increase with continued innovation and wider availability."[72]

At the time of writing (25 July 2024), the following English-language national statements were posted on the OPCW's EC-106 webpage that make reference to AI:[73]

> **Africa Group (Ghana)**: "The Africa Group takes note of the DG's note (S/2289/2024 dated 23rd May 2024) on Artificial Intelligence (AI) recognising the potential link between AI and chemical threats and its growing importance on the international disarmament agenda. The Africa group looks forward to

[71] OPCW 'Opening Statement by the Director-General to the 106th Session of the Executive Council' 106-DG.21 dated 6 July 2024 available at https://www.opcw.org/sites/default/files/documents/2024/07/ec106dg21%28e%29.pdf (accessed 18/07/2024)

[72] OPCW 'Note by the Director-General – Response to the Report of the Thirty Eighty Session of the Scientific Advisory Board' EC-106/DG.19 dated 4 July 2024 available at https://www.opcw.org/sites/default/files/documents/2024/07/ec106dg19%28e%29.pdf (accessed 18/07/2024)

[73] OPCW EC-106 documents available at https://www.opcw.org/resources/documents/executive-council/ec-106 (accessed 18/07/2024)

international meetings and activities as mentioned in the DG's note, including the global conference to be organised by the Kingdom of Morocco in collaboration with the Technical Secretariat on the role of AI in advancing implementation of the CWC in Rabat from 22nd – 24th October 2024."

In this regard, we congratulate the Kingdom of Morocco and the Technical Secretariat for organising a Global conference on the role of Artificial Intelligence in advancing the implementation of the CWC, scheduled on 22nd – 24th October 2024, in Rabat. We look forward to the discussions and outcomes that will provide a deeper understanding of the challenges and opportunities for solutions to these emerging threats as well as the potential opportunities for leveraging these new technologies for the benefit of the convention and the peaceful use of chemistry."

**Algeria** "The risk of chemical terrorism particularly through the potential use of chemical weapons by non-state actors, especially in zones of tension where they might be involved is aggravated by the rapid progress in science and technology, including artificial intelligence, which carries substantial potential to shape the development, deployment, and use of chemical weapons across various domains. We commend the DG's initiative in launching discussions on artificial intelligence with experts, on the potential impacts of AI on the implementation of the CWC and its evolving challenges. Even though, AI is a subject that still holds many undiscovered aspects, the early stage discussion have the merit of confirming the forward-thinking nature of our organization, which has contributed to its success."

**Australia**: "We should also be mindful of the risks posed by scientific and technological advancements. While offering potential benefits, emerging technologies - such as artificial intelligence and additive manufacturing - could pose serious risks to the OPCW's mission. If these technologies fall into the wrong hands, they may be harnessed in a way that gives rise to new threats, including the resurgence of chemical weapons"

**Bangladesh**: "While Artificial Intelligence applications in Chemistry may seem far away for countries like Bangladesh, it is actually not so given our footprint in global supply chains. We therefore look forward to the 'process' as well as 'outcome' of the Global Conference on AI and Chemistry in Morocco, this Fall, with much interest."

**Germany**: "It goes without saying that the discussions on these complex issues need to continue. This is why we are very much looking forward to another high-level conference in Rabat in October, organized jointly by Morocco and the TS, on the implications of AI for the CWC. Germany is confident that the Rabat conference will be able to build on the results of the Berlin conference, and has contributed 65.000 EUR to its realization.

Let me emphasize that Germany appreciates the early engagement of the TS under the leadership of the DG to address the challenges posed by emerging technologies and AI in particular. The workshop on AI in April, hosted by the DG at the CCT, produced important results that help to inform the continuing discussions at the OPCW on these complex issues. In this context, the contributions of the SAB will remain of critical importance. Having this in mind, we would like to suggest that the EC should have the opportunity to discuss the SAB's recommendations in a proper way instead of simply taking note of them.

It is clear that the topic of "emerging technologies" will remain high on the OPCW's agenda for the years to come. This is why we would deem it a good idea to put it as a regular item on the EC's agenda. "Emerging technologies" shows how important the input of and the exchange with external experts are to the work of the OPCW, be they scientists or industry representatives. With this in mind, the informal consultations, currently co-chaired by Germany and Costa Rica, have continued with the objective of making the ongoing interaction with civil society, including academia and chemical industry, even more meaningful and productive. We appreciate the engagement of States Parties, including many members of the Executive Council, and we call on them to continue constructively supporting this process."

**Japan**: "Developments in science and technology, particularly in the field of Artificial Intelligence (AI) are advancing at an unprecedented pace. AI has the potential to support the OPCW for our common goal of realizing a world where the possibility of the use of chemical weapons is completely excluded,

including preventing the re-emergence of chemical weapons. On the other hand, we should seriously consider the downside risk, which is that terrorists could invent new toxic chemicals using advanced technologies such as AI. In this regard, we support the initiative of DG Arias on holding events to gather insight from scientific experts on the current AI landscape and its prospects, while also fostering a deeper understanding of associated risks and opportunities. In addition, we welcome the Global Conference on the Role of Artificial Intelligence in Advancing the Implementation of the Convention in Morocco. We are positively considering to send participants to the conference to contribute to the discussion of the role of AI in the implementation of the Convention."

**Latvia**: "Rapid developments in science and technology, including Artificial Intelligence (AI), provide tremendous potential, but we must also acknowledge their ability to introduce new vulnerabilities that can be exploited by malicious actors (for instance, by abusing AI technologies to develop new chemical weapons). International cooperation and proactive approach are essential to addressing both the risks and opportunities posed by AI in the area of weapons of mass destruction. Latvia welcomes various initiatives in this respect – including a meeting organised by the Technical Secretariat with external experts and the RAC 2024 conference organised by Germany."

**Malaysia**: "Concerning Artificial Intelligence, as in the case of emerging technologies, it will have a mix of positive and negative impacts. Therefore, it is essential to conduct a thorough analysis to understand its potential implications for the OPCW's work and overall mission."

**Mexico**: "We welcome the document on artificial intelligence prepared by the Technical Secretariat and congratulate Germany and Morocco for convening conferences of international experts that will contribute to advancing the discussion of this important topic."

**Myanmar**: "Recognizing the transformative potential of artificial intelligence (AI) in enhancing the implementation and monitoring of the CWC, Myanmar expresses the readiness to participate in OPCW initiatives related to AI technology to explore and contribute to these advancements. We believe that AI technology can offer significant opportunities to strengthen the mechanisms for monitoring, verification, and compliance under the CWC. In this regard, we welcome the allocation of substantial financial, technical and human resources for developing countries from the OPCW in the realization of International Cooperation Activities to enable us not only to utilize the implications of AI technology but also to tackle its ethical, societal, and legal ramifications. We are also confident that the upcoming "Global Conference on the Role of Artificial Intelligence in Advancing the Implementation of the Convention" by the OPCW to be held in Rabat, Morocco, in October will be a complete success."

**North Macedonia**: "We also welcome the efforts initiated by the Director General to start looking at possible risks and opportunities that Artificial Intelligence (AI) could potentially bring to the implementation of the Chemical Weapons Convention, as well as deepening of Organisations' understanding of this fast developing technology."

**Pakistan**: "AI is a fast pacing important topic. We look forward to the discussion. As a multilateral institution, it would be important for OPCW to move on any topic including AI in an inclusive and transparent manner and staying loyal to its core mandate i.e. CWC."

**Poland**: "Poland thanks the Technical Secretariat and personally the DG for all efforts to meet challenges related to the emerging technologies, including the use of Artificial Intelligence (AI). We agree that this is a matter that needs to be closely monitored."

**Republic of Korea**: "Scientific and technological advances including Artificial Intelligence (AI) can bring valuable opportunities to effectively implement the Convention, but such advances could also be exploited for the potential re-emergence of the use of chemical weapons. In this context, the ROK welcomes the continued expansion and development of the ChemTech Center's activities since its inauguration in May 2023.

We also express support for the DG's initiatives on AI including the upcoming Global AI Conference which will be held in Morocco in October. This meeting will be a meaningful platform to discuss

opportunities and challenges AI can bring for the implementation of the Convention. I am pleased to announce that my government has decided to provide voluntary financial contribution to this event. Moreover, my delegation would like to encourage the TS to further strengthen its technical and professional capabilities to address current and future challenges associated with emerging disruptive technologies"

**Romania**: "From the perspective of the post-destruction phase, it is of paramount importance that the Secretariat deepens its understanding of Artificial Intelligence technology and explores its capabilities and opportunities in greater depth, but also the associated risks. To conclude, the development of technology must not make us forget the importance of engaging in the dialogue of the OPCW with all relevant stakeholders, ranging from the chemical industry or the scientific community to non-governmental organizations."

**Rwanda**: "Rwanda recognises the potential risk posed by artificial intelligence with regards to the use or threat of use of chemical weapons. We believe that a comprehensive approach between international organizations, State Parties and other relevant stakeholders is necessary in order to increase the preparedness to all eventualities.

**Slovenia**: "We appreciate the Director-General's efforts to explore the risks and potentials of new and emerging technologies, and in particular of artificial intelligence, to the implementation of CWC. It's important for the organization to remain fit for purpose, and to prepare for foreseeable challenges as much as possible."

**Switzerland**: "Particularly commendable are the efforts to leverage new and emerging technologies. In this context, we want to highlight the OPCW Artificial Intelligence Research Challenge. Additionally, we are looking forward to the outcomes of the upcoming Global Conference, convened by the Kingdom of Morocco and the Technical Secretariat, to explore and examine the implications of AI technology within the framework of the Convention. Switzerland has long been advocating for a robust and effective verification regime. We believe it is crucial to use our resources as efficiently as possible. In this sense, we are pleased that discussions on new and emerging technologies not only focus on potential risk but are also examining the potential benefits and applications to further enhance the work of the OPCW in all its aspects. The aforementioned conference will certainly make a valuable contribution in this regard.

**United Kingdom**: "We also note important DG led efforts to consider the implications of emerging technology for the OPCW. We look forward to further discussion of both the threats and opportunities presented by these new technologies."

**The EU's** statement under agenda item 18 noted that, inter alia, "The European Union welcomes the timely initiative by the Director-General in reaching out to available expertise to map out the ramifications of the AI and the associated risks and opportunities to the object and purpose of the Convention. We take note of the Director-General's note (S/2289/2024) on AI and OPCW Secretariat meeting with external experts, and its initial considerations. The EU believes this should be an inclusive process with active participation of all States Parties and input from external experts and civil society." Additionally, "the EU will co-finance a new AI Research Challenge as part of the EU's voluntary financial contribution in support of the OPCW activities."

**Germany** submitted a statement under agenda item 18 (AOB) which reflected on the Conference on Artificial Intelligence and Weapons of Mass Destruction that it hosted in Berlin on 28 June 2024. This statement noted, inter alia, that "[t]he panelists shared the view that AI does not act as a weapon by itself, but that its effect depended entirely on the intent of its use."

A joint Statement on behalf of the MIKTA countries (**Mexico**, **Indonesia**, **Republic of Korea**, **Republic of Türkiye**, and **Australia**) under agenda item 18 (AOB) commends efforts so far at looking at AI and notes, inter alia, "we look forward to advancing future policy discussions and supporting international cooperation and capacity building to identify and address opportunities and risks associated with new and emerging technologies. We will continue to advocate for collective responses to scientific and technological advancements that strengthen

agreed international commitments, and where the rules, norms, and standards on such developments are clear, mutually agreed, and consistently followed. To ensure the OPCW's work remains fit for purpose, we encourage the Technical Secretariat and States Parties to work with other international organisations and external stakeholders to explore additional ways to enhance our understanding of new and emerging technologies. In this light, and in addition to efforts to promote scientific literature, we would welcome further analysis of other technologies such as additive manufacturing, quantum computing, and biotechnology, and their implications for the Convention."

**12 July 2024**. The Organisation for the Prohibition of Chemical Weapons (OPCW) has announced that the Federal Republic of Germany has made a voluntary contribution of €65,000 to support the upcoming Global Conference on the Role of Artificial Intelligence (AI) in advancing the implementation of the Chemical Weapons Convention (CWC). The conference, scheduled to take place in Rabat, Morocco, from 22 to 24 October 2024, is designed to explore the implications of AI within the framework of the CWC.

This contribution was formalized on 12 July 2024 during a signing ceremony held at OPCW Headquarters in The Hague. The ceremony was attended by H.E. Mr. Thomas Schieb, Ambassador and Permanent Representative of Germany to the OPCW, and Ambassador Fernando Arias, OPCW Director-General. Germany's commitment to supporting the conference was emphasized by Ambassador Schieb, who noted,

> "We have no time to lose in trying to assess the effects of AI on the work of the OPCW. The discussions have only started: Building on the exchanges with high level experts at the conference on AI and Weapons of Mass Destruction in Berlin on 28 June and inspired by the programmatic keynote speech delivered by DG Arias in Berlin, the OPCW conference on AI in Rabat in October will carry the work forward. Germany thanks the Government of Morocco for hosting the conference and is happy to contribute financial support and technical expertise."

The conference in Rabat will address three key areas: the role and impact of AI in chemistry through evolving science and policy discourse, the challenges AI presents to the chemical industry, and the implications of AI in counterterrorism and the implementation of the CWC. The event is expected to bring together experts from science, industry, and government to facilitate a comprehensive understanding of AI's role in the context of the CWC.

Ambassador Arias, Director-General of the OPCW, acknowledged Germany's contribution as significant and timely, stating,

> "Germany's important contribution towards this timely conference on AI and the Chemical Weapons Convention is deeply appreciated. AI is a powerful tool that is transforming the chemical sciences. We must be prepared to address both the opportunities and challenges that it could present for the implementation of the CWC. This conference will foster dialogue among experts to ensure AI is used responsibly and contribute positively to global peace and security."

He also highlighted the OPCW's recent launch of an AI challenge, intended to explore how AI can strengthen the Organisation's capabilities and increase its readiness to address future challenges.

> "We encourage scientists and researchers from Member States to submit their proposals on how to use AI to enhance the OPCW's effectiveness, efficiency, and preparedness." [74]

[74] Organisation for the Prohibition of Chemical Weapons (OPCW). (2024). *Germany provides €65,000 to support OPCW conference on role of AI in chemical weapons*. Available at: https://opcwon.org/media-centre/news/2024/07/germany-provides-eu65000-support-opcw-conference-role-ai-chemical-weapons.html Accessed 24 Oct. 2024

**27 July 2024.** The New York Times recently published an article titled "A.I. May Save Us or May Construct Viruses to Kill Us," which explores both the potential benefits and risks of artificial intelligence (A.I.). This piece highlights growing concerns over A.I.'s dual-use nature, particularly in its application to biological weapons.

A key point raised in the article is the concern that A.I. could "dramatically reduce the barrier to entry" for non-experts to develop lethal pathogens. Jason Matheny, president of the RAND Corporation, underscores the gravity of the situation, noting that creating a virus capable of killing millions may now cost less than $100,000. While this figure illustrates the alarming accessibility of such technology, Matheny emphasizes that it costs far more to develop vaccines or antiviral treatments. The author highlights that A.I. not only accelerates advancements in fields like medicine but also opens doors for potentially catastrophic misuse.

The article outlines the challenges posed by synthetic biology, which may eliminate the need to steal existing viruses from labs. Citing historical examples, the article draws a parallel between past terrorist attacks, such as those carried out by Aum Shinrikyo in Tokyo in 1995, and the exponentially greater damage possible today with A.I. technology.

The piece emphasizes that while bioweapons pose a tangible threat, the risk is not confined to information access. The logistical challenge of physically generating pathogens and toxins remains a critical hurdle. Nonetheless, the rapid pace of A.I. development raises concerns that these barriers could erode over time, particularly with advances in synthetic biology and biotechnology. The potential for A.I. to aid in creating viruses tailored to specific races or individuals adds another disturbing dimension to the issue.

In its conclusion, the article stresses the need for continued research and robust governance to address these emerging threats. The author points out the balance between promoting innovation and ensuring national security, with experts like Susan Rice cautioning against leaving critical decisions in the hands of tech companies. The potential for A.I. to reshape global security dynamics, both positively and negatively, remains a focal point of ongoing discussions.[75]

**30 July 2024.** The National Telecommunications and Information Administration (NTIA) releases a report titled "Dual-Use Foundation Models with Widely Available Model Weights."[76] This document, mandated by Executive Order 14110, examines potential risks and benefits of open AI models, including implications for chemical and biological weapons (CBW).

The report identifies a key concern as AI models potentially "substantially lowering the barrier of entry for non-experts to design, synthesize, acquire, or use chemical biological, radiological, or nuclear (CBRN) weapons." It notes that while current models likely lack capabilities to significantly increase CBW risks, "the coming years could see fast and hard-to-predict changes in AI capabilities."

The report emphasizes that AI-related CBW risks extend beyond information access: "Information about how to design CBRN weapons may not be the highest barrier for developing them. Beyond computational design, pathogens, toxins, and chemical agents need to be physically generated, which requires expertise and lab equipment to create in the real world."

---

[75] New York Times. (2024). *A.I. May Save Us or May Construct Viruses to Kill Us*. Available at: https://www.nytimes.com/2024/07/27/opinion/ai-advances-risks.html?smid=nytcore-ios-share&referringSource=articleShare&tgrp=cnt&pvid=3E10520C-8740-496E-972A-51FFD010E368 [Accessed 24 Oct. 2024].

[76] NTIA 'Dual-Use Foundations Models with Widely Available Model Weights Report' Report, 30 July 2024 available at https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf (accessed 16/09/2024)

Highlighting uncertainties, the report states: "Further research is needed to properly address the marginal risk added by the accessibility and ease of distribution of open foundation models. For instance, the risk delta between jailbreaking future closed models for CBRN content and augmenting open models, as well as how the size of the model, type of system, and technical expertise of the actor, may change these calculations remains unclear."

The report also addresses concerns about AI's potential role in disinformation and misinformation. It notes that open foundation models could potentially "enable disinformation campaigns by adversarial actors." However, the report suggests that the impact of this risk may be limited, stating, "While many agree that open foundation models enable a larger range of adversarial actors to create disinformation, others dispute the importance of this assertion." The report cites arguments that "the bottleneck for successful disinformation operations is not the cost of creating it" and that "generative AI tools may prove useful to researchers and others combating disinformation." In relation to CBW, the report does not explicitly link disinformation to weapons development but includes it as part of the broader risk landscape associated with open AI models.

To address these and other risks, the report proposes a three-step framework: collecting evidence, evaluating that evidence against predefined thresholds, and acting on those evaluations if necessary. For CBW specifically, it recommends monitoring AI capabilities in synthesizing knowledge, advising on weapons production, and automating certain processes.

## AUGUST 2024

**7 August 2024.** Vox publishes an article by Jonas Sandbrink, a biosecurity researcher at the University of Oxford, titled "ChatGPT could make bioterrorism horrifyingly easy."[77] The piece examines how AI, particularly large language models like ChatGPT, could lower barriers to bioterrorism by providing easy access to specialized knowledge. Sandbrink argues that AI could help overcome historical bottlenecks in bioweapons development, citing examples like Aum Shinrikyo's failed 1990 botulinum toxin attacks. The article suggests AI could expand the pool of capable actors beyond state-level programs to include terrorist groups and individuals. It warns that biological design tools might enable the creation of pathogens "much worse than anything nature could produce." Sandbrink proposes mitigation strategies, including mandatory gene synthesis screening and pre-release evaluations of AI models. The article reflects growing concerns about AI's potential to democratize bioweapons knowledge while acknowledging its benefits in legitimate scientific research.

**22 August 2024** A recent *Science* article outlines both the benefits and risks of AI models applied to biological data. While these models have the potential to accelerate developments in drug and vaccine design and improve agricultural yields, the article underscores concerns about their possible misuse. The authors note that the same models used to create benign viral vectors for gene therapy could also be adapted to design harmful pathogens. They warn,

> "the same biological model able to design a benign viral vector to deliver gene therapy could be used to design a more pathogenic virus capable of evading vaccine-induced immunity."

Currently, biological AI models face significant limitations. The article describes them as producing "blurry image[s] of novel bacterial genomes" and notes that they are hampered by data availability and the need for in

[77] Sandbrink, J. 'ChatGPT could make bioterrorism horrifyingly easy' Vox, 7 August 2024 available at https://www.vox.com/future-perfect/23820331/chatgpt-bioterrorism-bioweapons-artificial-inteligence-openai-terrorism (accessed 16/09/2024)

vitro validation. Despite these constraints, the rapid development of AI technology suggests that future models may overcome these limitations, raising concerns about potential misuse.

Experts in the field, such as David Baker and George Church, have voiced concerns about protein-design technologies, highlighting that these innovations "are vulnerable to misuse and the production of dangerous biological agents." Similarly, the developers of genomic prediction models have acknowledged that AI tools could "catalyze the development of harmful synthetic microorganisms" if not properly managed. In response to these concerns, a global consortium of biological model developers has signed the *Responsible AI x Biodesign* statement. This agreement includes commitments to evaluate the risks associated with biological models before release and to adhere to biosecurity measures, such as sourcing synthetic nucleic acids from screened providers.

However, the *Science* article points out that voluntary commitments are insufficient. "The scientific community generally agrees that substantial risks require formal oversight processes," the article states, emphasizing that self-regulation is not enough. The authors suggest that governments, especially in the U.S. and U.K., should pass legislation to regulate advanced biological models. These regulations would focus on models requiring large computational resources or those trained on sensitive biological data to prevent the creation of pandemic-level pathogens. Both the U.S. and U.K. have begun establishing organizations to develop safety evaluations for advanced AI models. These evaluations aim to assess the risk posed by biological models and ensure that safeguards are in place before the models are made publicly available. The article argues that pre-release evaluations are necessary to mitigate the risks associated with AI-driven biological models. The report stresses that while current AI models may not yet pose substantial risks, advancements in data generation and AI's ability to manipulate biological sequences mean that the situation could change quickly. As the article notes, "the essential ingredients to create highly concerning advanced biological models may already exist or soon will." The article concludes with a call for balancing biosecurity with the need for scientific openness. Policymakers are urged to mitigate the risks of AI-driven biological models while allowing researchers the freedom to develop and share their findings. Regulations should be narrowly focused on models that pose significant risks, ensuring that the broader field of AI research remains unhindered.[78]

**27 August 2024.** Shopify publishes a blog post titled "Dangers of AI for Ecommerce: How To Mitigate Risks," discussing potential risks associated with AI in e-commerce.[79] The article briefly mentions biological weapons in the context of broader AI risks. In a section on "catastrophic risk," the post suggests that AI might be used in a global arms race to study natural pandemics and potentially create engineered pandemics as weapons. This reference to biological weapons is presented as an extreme scenario, separate from the main e-commerce focus of the article. It's used to illustrate a point about the potential dangers of AI without human values or moral constraints, rather than being directly relevant to the article's primary discussion of AI in e-commerce. The mention of biological weapons in this context demonstrates how concerns about AI's potential role in bioweapons development are permeating discussions even in fields not directly related to security or biosciences.

**29 August 2024.** The **Nucleic Acid Standards for Biosecurity Act** has been introduced to address biosecurity concerns surrounding the production of custom-ordered nucleic acids, such as DNA. With the increase in biotechnology production, there is a growing need for establishing industry standards that ensure nucleic acids

[78] **Doni Bloomfield, '**AI and biosecurity: The need for governance (2024) 385(6671) ***Science*** **doi:** 10.1126/science.adq1977 **accessed 29 September 2024**.

[79] Shopify 'Dangers of AI for Ecommerce: How to mitigate risks' Blog, 27 August 2024, available at https://www.shopify.com/uk/blog/dangers-of-ai# (accessed 16/09/2024)

are synthesized safely and securely. These standards aim to mitigate potential risks without hindering advancements in biotechnology, artificial intelligence (AI), or U.S. leadership in biomanufacturing.

The act would establish a consortium under the National Institute of Standards and Technology (NIST) to develop voluntary, consensus-based best practices and technical standards. This consortium will include representatives from industry, academia, nonprofit organizations, and other stakeholders. The goal is to create operational guidance to help the biotechnology industry and customers evaluate the performance of screening systems, ensuring biosecurity while promoting innovation.

Congresswoman Caraveo and Congressman McCormick, who co-sponsored the bill, emphasized the importance of balancing technological progress with security concerns.

"As a doctor, I'm a firm believer that we must embrace new technologies while developing standards and best practices to prevent threats of any kind," said Caraveo. Congressman McCormick highlighted the need to protect against biological threats, noting that "as Artificial Intelligence advances and accessibility grows, so do the risks."

Several key institutions have voiced their support for the legislation. Dr. Tom Inglesby, Director of the Johns Hopkins Center for Health Security, stated that the bill represents an important bipartisan step toward enhancing U.S. biomanufacturing while reducing biosecurity risks. Jason Green-Lowe, Executive Director of the Center for AI Policy (CAIP), explained that progress in AI could potentially allow malicious actors to design harmful biological compounds, making it essential to verify the safe manufacture of DNA. Eric Gastfriend, Executive Director of Americans for Responsible Innovation, added that the bill will help the U.S. establish leadership in developing global biosecurity standards.[80]

## SEPTEMBER 2024

**1 September 2024.** In *Pakistan Research Journal of Social Sciences*, Bushra Qamar publishes an article, *Risks of Bioterrorism Escalating Due to Artificial Intelligence*, examining how advancements in artificial intelligence (AI) may increase the accessibility and capabilities of bioterrorism. Qamar explores AI's dual-use potential, detailing how AI can assist in biosecurity efforts, while simultaneously introducing risks by lowering the technical barriers traditionally limiting bioweapon creation. The article emphasizes that AI tools, particularly large language models (LLMs) and biological design platforms, are reshaping access to biotechnological information that could be misused for harmful purposes.

Qamar's study points to LLMs' potential to support non-expert users in understanding complex laboratory procedures, synthesizing biological agents, and navigating pathogen design. AI systems also have the capacity to autonomously generate new biological compounds, including biotoxins, with fewer resources than required by conventional research. These advancements could, according to Qamar, shift the landscape of biosecurity by making bioweapon development more feasible for a wider array of actors.

The article calls for comprehensive global biosecurity measures, suggesting a regulatory framework for AI applications in the life sciences. This includes oversight of data access and control over AI-powered biological design tools to reduce risks associated with unauthorized use. Qamar concludes with an emphasis on balancing

[80] Rep Caraveo, 'Rep. Caraveo Introduces Bipartisan Legislation to Promote Nucleic Acid Biosecurity' (Caraveo.house.gov, 2023) https://caraveo.house.gov/media/press-releases/rep-caraveo-introduces-bipartisan-legislation-promote-nucleic-acid-biosecurity accessed 29 September 2024.

innovation with security, recommending that international stakeholders collaborate to develop safeguards against the unintended proliferation of AI-enabled biotechnological knowledge.[81]

**5 September 2024.** The Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (CETS No. 225) was opened for signatures during a conference of Council of Europe Ministers of Justice in Vilnius. This treaty represents the first legally binding international framework aimed at ensuring AI technologies comply with established standards for human rights, democracy, and the rule of law. As of the date of writing, the framework Convention was signed by several Council of Europe member states, including Andorra, Georgia, Iceland, Norway, the Republic of Moldova, San Marino, the United Kingdom, the European Union and non-European signatories such as the United States and Israel. The Convention aims to create a legal framework that ensures AI systems are developed and used in a manner consistent with human rights. Council of Europe Secretary General Marija Pejčinović Burić commented on the purpose of the treaty:

> *"The Framework Convention is designed to ensure that AI upholds our standards. It is a text developed through an open and inclusive approach, benefiting from multiple expert perspectives."*

The Framework Convention provides a legal structure that covers the entire lifecycle of AI systems, from development to deployment and decommissioning. It focuses on promoting innovation in AI technologies while managing potential risks to human dignity, equality, and individual rights. According to the Convention, signatories must adopt "appropriate legislative, administrative, or other measures" to address potential impacts AI might have on human rights, democracy, and the rule of law.[82]

To promote transparency and accountability within the realm of AI development, signatories are required to establish oversight mechanisms to monitor AI's impact on human rights and democracy. This includes ensuring that AI systems do not worsen inequalities or discrimination, particularly in digital contexts. The framework also emphasizes public awareness when interacting with AI-generated content and ensures transparency in how these systems influence public debate.[83]

The convention was adopted by the Council of Europe's Committee of Ministers on 17 May 2024, and involved negotiations among 46 Council of Europe member states, the European Union, and 11 non-member states, including Argentina, Australia, Canada, and Japan. Representatives from the private sector, civil society, and academia participated as observers during the drafting process. Marija Pejčinović Burić added:

> *"I hope that these signatures will be followed by ratifications so that the treaty can enter into force as soon as possible."*

The treaty will come into force once five signatories, including at least three Council of Europe member states, have ratified it. Countries from across the globe will be eligible to join the treaty and comply with its provisions.

The Framework Convention takes a neutral approach, aiming to balance AI innovation with necessary protections for human rights. It encourages "safe innovation" to promote AI technologies while ensuring they do not negatively impact democratic values or personal freedoms. The Convention also focuses on digital literacy, urging member states to promote education on AI systems to help citizens navigate these technologies safely. The Council of Europe's Framework Convention on Artificial Intelligence establishes a legal framework designed

---

[81] Qamar, B. 'Risks of Bioterrorism Escalating Due to Artificial Intelligence' (*Pakistan Research Journal of Social Sciences*, 2024) 3(3) available at https://prjss.com/index.php/prjss/article/view/145 (accessed 11/11/2024)

[82] Council of Europe. (2024). *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225)*. Available at: https://www.coe.int (Accessed 6 Oct. 2024)

[83] Council of Europe. (2024). *Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225)*. Available at: https://www.coe.int (Accessed 6 Oct. 2024)

to regulate AI technologies in accordance with human rights standards, democracy, and the rule of law. As signatories begin the ratification process, the treaty provides an international standard for AI governance, aimed at addressing the challenges and opportunities posed by rapid technological advancements in AI.

**10 September 2024.** In the *International Journal for Conventional and Non-Conventional Warfare*, Jannat Naseeb publishes *Biological and Chemical Weapons: Contemporary Threats and Responses*, examining the modern landscape of biological and chemical weapons (BCW) threats and evaluating international responses. Naseeb identifies key drivers for BCW interest among state and non-state actors, citing motivations such as cost-efficiency, operational ease, and asymmetric warfare advantages. Through examples like Syria's use of sarin and chlorine gas and the 2001 anthrax attacks, Naseeb demonstrates how BCW can be deployed to inflict strategic and psychological damage, complicating regional stability and security.

The article emphasizes that advancements in artificial intelligence (AI) and biotechnology are intensifying these threats. Naseeb notes that AI, alongside synthetic biology and bioinformatics, is increasingly accessible and enables the rapid development and modification of pathogens. AI-driven tools can streamline complex bioengineering processes, allowing actors with limited resources to manipulate biological agents with heightened transmission or lethality potential. AI's capability to automate and simplify these processes, he argues, lowers traditional barriers to weapon development and exacerbates the dual-use dilemma of biotechnological innovations.

Naseeb critiques the limitations of existing international frameworks, such as the Biological Weapons Convention (BWC) and Chemical Weapons Convention (CWC), which, while comprehensive, face enforcement and verification challenges. He references the Organization for the Prohibition of Chemical Weapons (OPCW) findings on non-compliance cases, especially in Syria, to underscore gaps in regulatory enforcement. Naseeb advocates for enhanced global cooperation and adaptive regulatory frameworks to respond to the evolving risks AI presents, suggesting the development of "dynamic" oversight mechanisms tailored to rapid technological advancements.

To strengthen BCW countermeasures, the article recommends international collaboration on AI-informed verification technologies and the integration of real-time biosensors and molecular detection tools into monitoring systems. Naseeb concludes that, given AI's transformative potential in biotechnology, proactive policy and regulatory adaptation are essential to mitigate misuse risks and maintain effective BCW deterrence.[84]

**10 September 2024.** Bloomberg's Editorial Board published an article titled "Could AI Create Deadly Biological Weapons? Let's Not Find Out," discussing the potential risks of AI in facilitating the development of biological weapons.

The article frames AI, particularly large language models and biological design tools, as potentially lowering barriers to creating biological weapons by synthesizing knowledge rapidly and speeding up pathogen development. It cites an experiment where a chatbot advised MIT students on engineering deadly pathogens within an hour.

The editorial emphasizes that while current barriers to producing workable bioweapons remain high, technological advancements could change this in a few years. The article presents a range of potential threats, including rogue scientists, states, and terrorists, as well as acknowledging ongoing bioweapons programs in countries like Russia and North Korea.

[84] Naseeb, J. 'Biological and Chemical Weapons: Contemporary Threats and Responses' (*International Journal for Conventional and Non-Conventional Warfare*, 2024) 1(3) available at https://www.researchcorridor.org/index.php/ijcnw/article/view/77 (accessed 11/11/24)

The article suggests several preventive measures:

1. Government screening of powerful AI models, especially those trained on biological data
2. Stricter oversight of synthetic nucleic acid providers and DNA synthesis equipment
3. International cooperation on biosecurity best practices
4. Strengthening pandemic defenses and public health systems

The framing presents AI as a dual-use technology, potentially aiding both the creation and detection of new pathogens, as well as vaccine development. It portrays the primary risk as the democratization of bioweapons expertise, rather than the creation of entirely new types of weapons.

The article suggests that AI could lower barriers to bioweapons development for a range of actors, potentially making such weapons more accessible and attractive to a wider array of entities. It emphasizes the need for broader safeguards and improved defenses against potential biological threats, regardless of their source.[85]

**8 – 11 September 2024.** Switzerland's Spiez Laboratory held its bi-annual CONVERGENCE Conference, this year focusing on five main topics. These were precise editing in chemistry and biology; digitalisation, automation and artificial intelligence; manufacturing chemicals; therapeutic applications and drug delivery; and threat-agnostic biodefense. The conference report will be available at the end of 2024.[86]

**11 September 2024.** OPCW Director-General Fernando Arias visits the United Kingdom for high-level meetings, including discussions with the Minister of State for Energy Security and Net Zero, Lord Hunt of Kings Heath OBE.[87] The talks covered the global security landscape, situations in Syria and Ukraine, and recent advancements in science and technology. Particular attention was given to the implications of artificial intelligence on chemical sciences for enhancing capabilities in chemical weapons monitoring and verification. Director-General Arias highlighted the UK's role in co-funding the OPCW Artificial Intelligence Research Challenge, stating, "I am pleased to highlight the UK's role in co-funding the OPCW Artificial Intelligence Research Challenge, an initiative that supports researchers and scientists in developing innovative AI solutions that can enhance the OPCW's effectiveness, efficiency, and preparedness in implementing the Chemical Weapons Convention."

**12 September 2024.** OpenAI announces they are developing a new series of AI models, with the first in series named OpenAI 01, that can "reason through complex tasks and solve harder problems than previous models in science, coding, and math."[88] The blog entry contains a link to their system card, a report presenting the models capabilities and apparent risks, with risks being categorised as low, medium, high, and critical.[89]

According to OpenAI's evaluations, the 01 models demonstrated superior performance in biological threat information tasks, outperforming previous models and sometimes even human experts. The system card noted that "o1-preview and o1-mini both performed well on each of the four physical success biothreat information questions (Acquisition, Magnification, Formulation, and Release), often achieving significant uplifts over GPT-4o." Experts who tested the 01-preview model found it "significantly useful for answering questions beyond

---

[85] Brandon Judkins, 'Could AI Create Deadly Biological Weapons? Let's Not Find Out' (GMToday, 2024) https://www.gmtoday.com/daily_news/commentary/could-ai-create-deadly-biological-weapons-let-s-not-find-out/article_bb2c84d4-7814-11ef-bf1a-7b66e8cbce08.html accessed 29 September 2024.

[86] Spiez Laboratory. (2024.). *International Conferences at Spiez Laboratory*. Available at: https://www.spiezlab.admin.ch/en/internationale-konferenzen-im-labor-spiez-en [Accessed 24 Oct. 2024]

[87] OPCW 'OPCW Director-General visits UK, meetings with Minister of State for Energy Security, other high officials, to discuss global security and technological advances' News 11 September 2024 available at https://www.opcw.org/media-centre/news/2024/09/opcw-director-general-visits-uk-meets-minister-state-energy-security (accessed 16/09/2024)

[88] OpenAI 'Introducing OpenAI 01-preview" 12 September 2024 available at https://openai.com/index/introducing-openai-o1-preview/ (accessed 16/09/2024)

[89] OpenAI 'OpenAI 01 System Card' 12 September 2024 available at https://cdn.openai.com/o1-system-card.pdf (accessed 16/09/2024)

access to the internet," with particular praise for its ability to speed up research processes and provide detailed protocols. The models also showed improved capabilities in tacit knowledge and troubleshooting related to biological experiments.

However, OpenAI was careful to highlight the models' limitations. They emphasized that "the models do not enable non-experts to create biological threats, because creating such a threat requires hands-on laboratory skills that the models cannot replace." The 01 models struggled with fully automating wet lab work and couldn't independently design DNA constructs without external tools.

OpenAI classified both 01-preview and 01-mini as "medium risk" for biological threat creation. The system card stated, "Our evaluations found that o1-preview and o1-mini can help experts with the operational planning of reproducing a known biological threat, which meets our medium risk threshold."

A concerning aspect noted in the system card was the models' potential to provide incomplete safety information. In some cases, they gave "very misleading safety information omitting things like toxic byproducts, explosive hazards, carcinogens, or solvents that melt glassware."

The release of the 01 models and their associated system card highlighted the dual-use nature of advanced AI in biological research. While these models promised to accelerate beneficial research, they also raised concerns about potential misuse. The development underscored the growing need for robust safety protocols, ethical guidelines, and careful consideration of access controls in the rapidly evolving field of AI-assisted biological research.

**12 September 2024.** The Bulletin of the Atomic Scientists published an article titled "Apathy and hyperbole cloud the real risks of AI bioweapons", written by Filippa Lentzos, Jez Littlewood, Hailey Wingo, and Alberto Muti.[90] Addressing growing concerns about artificial intelligence's potential role in biological weapons development, the article reflects on how AI, particularly large language models, might lower barriers to accessing dual-use knowledge while noting that historical evidence suggests biological weapons development remains complex. They cautioned against both dismissing and exaggerating AI's impact on biosecurity. The article called for structured risk assessments and emphasized the need to consider AI within the broader context of emerging technologies. It highlighted the importance of developing collaborative responses at various levels to address evolving biosecurity challenges. This publication contributed to ongoing discussions about balancing technological advancement with biosecurity concerns in an era of rapid AI development.

**13 September 2024.** The Carnegie Endowment for International Peace published a report by Holden Karnofsky titled "If-Then Commitments for AI Risk Reduction."[91] The report discusses the potential use of if-then commitments to mitigate risks associated with artificial intelligence, including those related to chemical and biological weapons (CBW). Karnofsky presents a framework where AI developers commit to implementing specific risk mitigations if their AI models reach certain capability thresholds. For CBW, the author proposes a "tripwire" capability: an AI model's ability to interactively advise a malicious actor to the point where they would have a substantial chance of successfully producing and releasing a catastrophically damaging CBRN weapon.

The report frames AI-assisted CBW development as a risk of proliferation, where AI could serve as a virtual substitute for expert advisers, potentially expanding the number of people capable of producing and deploying such weapons. The article discusses AI's potential role in lowering barriers to entry for non-experts in bioweapons development. While it doesn't explicitly mention creating new types of weapons, it does suggest that AI could potentially make existing bioweapons more accessible to a wider range of actors. Karnofsky's

[90] Lentzos, F. et al 'Apathy and hyperbole cloud the real risks of AI bioweapons' The Bulletin of the Atomic Scientists, 12 September 2024 available at https://thebulletin.org/2024/09/apathy-and-hyperbole-cloud-the-real-risks-of-ai-bioweapons/ (accessed 16/09/2024)

[91] Karnofsky, H. 'If-then commitments for AI risk reduction' Carnegie, 13 September 2024 available at https://carnegieendowment.org/research/2024/09/if-then-commitments-for-ai-risk-reduction?lang=en (accessed 16/09/2024)

approach assumes that only a small percentage of the population currently has the expertise to develop CBW, and even fewer would want to. The framework implies that the primary risk is from determined individuals or terrorist organizations gaining access to expert-level knowledge through AI, rather than state actors or those already possessing expertise.

The report suggests that the utility of CBW remains limited, but AI could potentially make them more attractive to a wider range of actors by reducing the expertise barrier. This framing focuses on the informational aspect of AI assistance rather than physical production capabilities.

**14 September 2024.** Vox reports on some of the risks reported in OpenAI's latest model, 01 (also nicknamed Strawberry), noting,

> Here are some things that are not cool: Nuclear weapons. Biological weapons. Chemical weapons. And according to OpenAI's evaluations, Strawberry can help people with knowledge in those fields make these weapons. [...] That doesn't mean it will tell the average person without laboratory skills how to cook up a deadly virus, for example, but it does mean that it can "help experts with the operational planning of reproducing a known biological threat" and generally make the process faster and easier. Until now, the company has never given that medium rating to a product's chemical, biological, and nuclear risks."[92]

**14 September 2024.** Paul Scharre, vice president and director of studies at the Center for a New American Security, publishes an article in Time titled "Regulating AI Is Easier Than You Think". The article discusses approaches to governing AI technology, drawing parallels with nuclear technology regulation. Key points relevant to chemical and biological weapons include:

1. Scharre notes that AI, as a general-purpose technology, could be used to develop chemical or biological weapons, highlighting the dual-use nature of AI.
2. The article proposes regulating AI by controlling access to specialized chips needed to train advanced AI models, similar to how nuclear technology is regulated by controlling access to fissile materials.
3. Scharre suggests that this approach could prevent "adversary nations, terrorists, or criminals from using the most powerful AI systems," which could have implications for CBW development.
4. The article mentions that the U.S. government has begun debating restrictions on the most powerful trained AI models and how widely they can be shared, which could affect potential misuse for CBW purposes.
5. Scharre proposes an international governance framework for AI, similar to nuclear non-proliferation efforts, which could potentially address CBW-related concerns.

The article does not focus specifically on chemical and biological weapons but presents a broader argument for AI governance that could have implications for preventing AI misuse in CBW development.[93]

**17 September 2024.** The Nuclear Threat Initiative (NTI) | bio and the Center for Arms Control and Non-Proliferation (CACNP) host a Congressional staff briefing titled "Preventing Biological Catastrophe and Protecting the U.S. Bioeconomy." The briefing, reported on 26 September, focuses on the convergence of

[92] Samuel, S. 'The new followup to ChatGPT is scarily good at deception' Vox, 14 September 2024 available at https://www.vox.com/future-perfect/371827/openai-chatgpt-artificial-intelligence-ai-risk-strawberry (accessed 16/09/2024)

[93] Scharre, P. 'Regulating AI Is Easier Than You Think' TIME 14 September 2024 available at https://time.com/7021171/ai-regulation-chips/ (accessed 26/09/24)

artificial intelligence and biotechnology, and the associated risks and governance challenges. Key points include:

1. Discussion of rapidly advancing capabilities arising from the combination of AI and biotechnology.
2. Exploration of oversight measures for dual-use life science research of concern.
3. Consideration of establishing a new agency to lead innovation while integrating biosecurity and biosafety into life sciences research and biotechnology.
4. Proposal for a legal requirement to screen DNA synthesis orders, coupled with implementation incentives.

The briefing reflects growing concern about the potential misuse of AI in biological research and the need for proactive governance measures to mitigate risks while fostering innovation in the bioeconomy.[94]

**17 – 18 September 2024.** The Centre for the Study of Existential Risk (CSER) hosts the Cambridge Conference on Catastrophic Risk 2024 in Cambridge, UK. This interdisciplinary conference aims to bring together students, researchers, practitioners, and policymakers to discuss global catastrophic risks (GCRs). The event features sessions on various topics, including the intersection of systemic risk and GCR, historical perspectives on future crises, emerging and converging risks, disaster risk reduction for GCRs, and intergovernmental governance of catastrophic risks. Of particular note is a session on "Emerging and Converging Risks," which invites submissions on "evidence-based assessments of areas of convergence between risks, such as AI and biological risk." This inclusion highlights interest in the potential interactions between artificial intelligence and biological risks, in relation to global catastrophic threats.[95]

**19 September 2024**. France contributes €125,000 to the OPCW, allocated as follows: €55,000 for a project on preventing illicit chemical transfers in French-speaking African countries; €50,000 to strengthen chemical safety legislation in Central Asia, Sub-Saharan Africa, and the Caribbean; and €20,000 to support the upcoming conference on Artificial Intelligence and chemical weapons in Rabat. A news item, published on 23 September, notes that this voluntary contribution aligns with France's priorities in supporting OPCW activities, particularly in French-speaking Africa, chemical safety, and addressing terrorist threats.[96]

**19 September 2024.** As reported on the English version of China News Service, Chinese Minister of Industry and Information Technology Jin Zhuanglong met with OPCW Director-General Fernando Arias in Beijing. The meeting covered topics including industrial verification regime reform and the impact of new technologies on the chemical industry. Arias expressed hope to enhance cooperation with China “to address the impacts of complex international situations, chemical industry developments, and new technologies such as chemistry and artificial intelligence on compliance of OPCW.”[97]

---

[94] ‘NTI | bio Briefs Capitol Hill on Innovative Solutions to Biotechnology Risks’ NTI News item 26 September 2024 available at https://www.nti.org/news/nti-bio-briefs-capitol-hill-on-innovative-solutions-to-biotechnology-risks/ (accessed 30/09/2024)

[95] CSER ‘Cambridge Conference on Catastrophic Risk 2024’ available at https://www.cser.ac.uk/events/cccr-2024/ (accessed 19/09.2024)

[96] French Embassy, The Hague ‘OPCW - New French voluntary contribution to support the activities of the Organization for the Prohibition of Chemical Weapons (The Hague, 23 September 2024)’ News, 24 September 2024 available at https://nl.ambafrance.org/OPCW-New-French-voluntary-contribution-to-support-the-activities-of-the-27724 (accessed 26/09/2024)

[97] ECNS ‘China reiterates willingness to strengthen cooperation with OPCW’ ECNS Wire 19 September 2024 available at http://www.ecns.cn/news/cns-wire/2024-09-19/detail-ihehhmat3658543.shtml (accessed 26/09/24)

**19 September 2024.** Nicholas Emery-Xu, Richard Jordan, and Robert Trager publish "International governance of advancing artificial intelligence" in AI & SOCIETY. The paper discusses various approaches to governing transformative AI (TAI) technologies, with several significant mentions of chemical and biological weapons (CBW). Key points related to CBW include:

1. The paper opens the example of AI's potential in CBW development:

"Researchers at Collaborations Pharmaceuticals, a small drug company in Raleigh, NC, used artificial intelligence (AI) techniques to search for toxic molecules. After a few hours, they found 40,000 potential toxins. Some were known toxins, like the nerve agent VX, the most toxic chemical yet discovered, but many were predicted to be orders of magnitude more toxic than VX."

This example is used to highlight AI's capability to rapidly accelerate the discovery of potentially harmful substances.

2. The authors discuss the dual-use nature of AI, emphasizing its potential applications in both beneficial research and weapons development:

"Future models built on these foundations will automate aspects of R&D processes across numerous military and civilian domains, rapidly uncovering wonders and new technologies of destruction. The 40,000 toxins of Collaborations Pharmaceuticals are the very tip of this iceberg, and no one knows where these developments will lead."

3. When discussing the Biological Weapons Convention as an example of international governance, the paper notes its limitations:

"The Biological Weapons Convention, for instance, which major powers signed, limits the 'development, production and stockpiling' of biological weapons. Yet, this was not a significant military technology without military substitutes: biological weapons' utility was limited because of the potential harm to one's own side, and nuclear weapons offered a clear (and superior) substitute technology."

The authors further point out that the Convention lacks verification provisions and has been "flagrantly violated by the Soviet Union and others."

4. The paper frames the potential for AI to lower barriers for non-state actors to develop CBW as a significant concern:

"If AI uncovers destruction-dominant technologies, a single sociopath could wreak such damage as to make some control of the technology a matter of life-and-death."

In their conclusion, the authors acknowledge that no single governance approach is likely to be sufficient for addressing the complex challenges posed by AI, including its potential misuse in CBW development. They suggest that a combination of strategies may be necessary, potentially including elements of non-proliferation regimes, verifiable limits, and international monopolies. The paper emphasizes the urgency of developing governance mechanisms, stating, "If the technology evolves quickly, we are about to live through unprecedented times; it should not surprise us if unprecedented political structures emerge—though we should not blithely assume they will."

These explicit mentions of CBW are used to illustrate the broader challenges of governing transformative AI technologies. The paper emphasizes the need for robust international governance to mitigate these risks while acknowledging the difficulties in implementing effective control measures for dual-use technologies like AI.[98]

[98] Emery-Xu, N. et al 'International governance of advancing artificial intelligence' (*AI &SOCIETY*, 2024) available at https://link.springer.com/article/10.1007/s00146-024-02050-7 (accessed 26/09/2024)

**21 September 2024**. Decrypt reports on testimony given to the U.S. Senate Committee on the Judiciary Subcommittee on Privacy, Technology, & the Law by William Saunders, a former OpenAI employee. The testimony highlights growing concerns about AI's potential to aid in the development of biological weapons:

1. OpenAI's newest GPT-o1 model is described as "the first system to show steps towards biological weapons risk," capable of assisting experts in planning to reproduce known biological threats.
2. The potential for "catastrophic harm" is emphasized if AGI (Artificial General Intelligence) systems are developed without proper safeguards.
3. Saunders warns that an AGI system could potentially be built within three years, raising urgent concerns about AI safety and governance.
4. The article frames the biological weapons risk as part of a broader set of concerns about AGI, including potential for misuse, lack of adequate safety measures, and prioritization of deployment over rigorous safety testing.
5. Multiple high-profile departures from OpenAI are mentioned, with several former employees and board members citing concerns about the company's approach to AI safety.

The narrative implicates both AI companies (particularly OpenAI) and potential malicious users as sources of risk. The framing suggests a rapid acceleration of AI capabilities that outpaces safety measures and regulatory frameworks, and Directly links cutting-edge AI models to biological weapons risks.

- Suggests that even legitimate AI research could inadvertently enable CBW development.
- Highlights the challenges of governing dual-use AI technologies.
- Indicates growing concern among AI insiders about safety and security implications of advanced AI systems.
- Demonstrates how CBW risks are being used to argue for urgent regulatory action in AI development.[99]

**22 September 2024.** Arkansas City, Kansas, reports a cybersecurity incident at its Water Treatment Facility.[100] While the attack does not explicitly mention AI or chemical/biological weapons, it highlights the growing vulnerability of critical water infrastructure to cyber threats, which could potentially be exploited for CBW attacks.

City Manager Randy Frazer stated, "Despite the incident, the water supply remains completely safe, and there has been no disruption to service." The facility switched to manual operations as a precautionary measure. This incident follows a trend of increasing cyberattacks on U.S. water utilities, with the EPA reporting that about 70% of inspected utilities violated cybersecurity standards in the past year.

The article notes that foreign cybercriminal groups, particularly those affiliated with Russia and Iran, have been targeting smaller, more vulnerable water utilities in rural areas. While not explicitly mentioning AI, these sophisticated attacks could potentially involve AI-driven tools or techniques.

---

[99] Lanz, J. 'OpenAI's new AI shows "steps towards biological weapons risks", ex-staffer warns senate' Decrypt 21 September 2024 available at https://decrypt.co/250568/opena-new-ai-steps-towards-biological-weapons-risks-warns-senate (accessed 26/09/2024)

[100] Jain, S. 'Arkansas City Responds to Cybersecurity Incident at Water Facility, Ensures Safe Drinking Water' The Cyber Express 25 September 2024 available at https://thecyberexpress.com/arkansas-city-water-treatment-facility-attack/ (accessed 30/09/2024)

This event underscores the potential for cyber vulnerabilities in water treatment facilities to be exploited for biological attacks, even if not realized in this specific incident.

**23 September 2024.** The Foreign Policy Research Institute publishes an article by Mohammed Soliman and Vincent J. Carchidi titled "Re-Balancing the Strategy of Tech Containment."[101] The piece discusses the U.S. strategy of technological containment against China, with a focus on export controls and their implications for advanced technologies, including AI.

While the article does not extensively discuss chemical and biological weapons in relation to AI, it does mention them in the context of emerging legislative efforts:

"Notably, the House Foreign Affairs Committee passed the Enhancing National Frameworks for Overseas Critical Exports Act by a 43-3 vote in May. The legislation aims to amend the Export Control Reform Act of 2018 in two main ways: first, by clearly empowering the US President to restrict the export of AI systems (as opposed to merely their enabling hardware), and second, by restricting Americans from working with foreigners to develop AI systems deemed a risk to national security."

The authors note that this legislation aligns with executive branch policymaking on AI, referencing President Biden's October 2023 Executive Order:

"The legislation extends a theme in Executive Branch AI policymaking in which some AI models—current or future—are believed to potentially lower the barrier of entry for malicious actors to design, build, acquire, or use chemical, biological, radiological, or nuclear weapons of mass destruction."

The article frames these developments as part of a broader U.S. strategy to maintain technological superiority over China, particularly in AI and advanced computing. The authors argue for a more balanced approach to export controls, suggesting that overly broad restrictions could have unintended consequences, including pushing potential partners closer to China and undermining U.S. technological leverage in the long term.

The authors recommend a two-pillar approach: 1) a proactive export control regime targeting the most sensitive technologies, and 2) establishing a technology dialogue with allies to coordinate on new controls. This approach aims to balance the need for technological containment with maintaining strong international partnerships and avoiding overreach in restrictions.

**24 September 2024**. Mariam Elgabry and Shane Johnson publish "Cyber-biological convergence: a systematic review and future outlook" in Frontiers in Bioengineering and Biotechnology. The paper examines the intersection of engineered biology and cybersecurity, highlighting both opportunities and risks as biological systems become increasingly integrated with digital technologies.

The authors identify several key areas where AI is transforming biological research and manufacturing, including automated bio-foundries, the Internet of Biological Things, and AI-driven drug discovery. However, they also emphasize the potential for misuse, particularly in the context of chemical and biological weapons.

When discussing CBW, the paper focuses primarily on biological agents, particularly engineered pathogens. The authors highlight a concerning example where MIT students used large language models to design potential pandemic pathogens in just one hour. They also reference a case where an AI-powered

[101] Soliman, M. and Carchidi, V. 'Re-Balancing the Strategy of Tech Containment' FPRI 23 September 2024 available at https://www.fpri.org/article/2024/09/re-balancing-the-strategy-of-tech-containment/ (accessed 30/09/2024)

drug discovery algorithm was inverted to design 40,000 potential biochemical weapons in under 6 hours. These examples underscore the dual-use nature of AI in biology and the potential for rapid development of dangerous agents.

The paper considers a range of actors who might exploit these technologies, from individual "biohackers" and terrorist groups to state-sponsored actors. The authors note that the democratization of biotechnology, combined with AI assistance, could lower barriers to entry for developing biological weapons, making them accessible to a wider range of malicious actors.

The study frames the convergence of cyber and biological technologies as a double-edged sword, offering significant benefits for medicine and agriculture while simultaneously creating new vulnerabilities. The authors highlight emerging threats such as bio-malware, neuro-hacking, and the potential for AI-enhanced biological attacks. They specifically mention the risk of engineered pathogens designed to mimic common diseases, potentially camouflaging their initial spread.

To address these risks, the paper calls for the development of a new discipline: cyberbiosecurity. This field would encompass traditional biosafety and biosecurity measures while also addressing the unique challenges posed by the integration of biological and digital systems. The authors suggest several policy recommendations, including the need for standardized data exchange formats, improved cyber hygiene practices, and the development of adversary-resilient biological protocols.

A key takeaway is the urgent need for proactive governance and security measures in the rapidly evolving field of engineered biology. The paper emphasizes that as AI continues to advance, it's crucial to develop robust safeguards and ethical frameworks to prevent misuse while still fostering innovation. The authors conclude by calling for more inclusive research studies and greater public participation in biosecurity efforts, recognizing that the challenges posed by cyber-biological convergence require a multidisciplinary and collaborative approach.[102]

**25 September 2024.** The Conversation publishes an article by Shweta Singh, Assistant Professor at Warwick Business School, discussing OpenAI's new AI system called Strawberry (also known as o1). The article reports that OpenAI's own evaluations rate Strawberry as a "medium risk" for its ability to assist experts in the "operational planning of reproducing a known biological threat." Singh argues that OpenAI's policy of allowing "medium risk" models to be released for wide use underestimates the potential threat, particularly if manipulated by bad actors. The article highlights ongoing concerns about AI's potential role in biological weapons development and calls for stronger regulatory frameworks and scrutiny protocols for AI models like Strawberry.[103]

**25 September 2024.** Stanford Law School publishes an interview with Professor Mark Lemley discussing California's AI Safeguards Bill (SB 1047). Lemley states that "SB 1047 is designed to reduce the risk of catastrophic 'rogue AI' by requiring AI companies to evaluate all AI projects for a risk of nuclear or chemical war or other catastrophes and to code in an 'off switch' that could shut down AI that went rogue." He expresses skepticism about the risk of "rogue AI," arguing that "the risk of rogue AI is significantly overstated." Lemley notes that while "there are certainly AI projects that pose real risks, like autonomous weapons systems," applying these safeguards broadly "seems like overkill." The interview does not specifically address biological weapons, focusing instead on the broader implications of the bill for AI

[102] Elgabry M and Johnson S 'Cyber-biological convergence: a systematic review and future outlook' (*Front. Bioeng. Biotechnol*, 2024) 12

[103] Singh, S. 'OpenAI's Strawberry program is reportedly capable of reasoning. It might be able to deceive humans' The Conversation 25 September 2024 available at https://theconversation.com/openais-strawberry-program-is-reportedly-capable-of-reasoning-it-might-be-able-to-deceive-humans-239748 (accessed 26/09/2024)

development and regulation. Lemley argues that "regulation should be national (or ideally global)," but acknowledges that congressional gridlock has led to state-level initiatives.[104]

**25 September 2024.** Yaniv Golan publishes an article on Medium titled "Assessing AI Risks: AI Bets 50% on AI-induced Catastrophe Within 10 Years". The article uses commercially available AI language models (ChatGPT 4o and ChatGPT o1-preview) to assess the risk of AI causing a major catastrophic event within the next decade. While not specifically focused on chemical and biological weapons, the analysis touches on related concerns:

1. The potential for AI to assist in bioweapon design is noted as a near-term concern, with researchers demonstrating AI's ability to design toxic molecules.
2. The article references a 2022 incident where scientists modified an AI system to generate 40,000 candidate molecules for chemical weapons in six hours.
3. The analysis suggests AI could lower barriers for non-state actors to develop chemical or biological weapons by providing expert-level knowledge.
4. Concerns are raised about AI integration into critical infrastructure, including nuclear command and control systems, which could increase risks of unintended escalation.

The AI models estimate a 50% likelihood of an AI-induced catastrophic event within 10 years. The article's methodology involves using AI models as "odds makers" and iteratively refining estimates based on information gathered through an AI-powered search engine.

However, the article acknowledges several limitations and potential biases in its approach, as outlined in an appendix featuring criticism from another AI model (Claude 3.5):

1. Potential overreliance on sources emphasizing AI risks.
2. Lack of historical context on managing other transformative technologies.
3. Overemphasis on speculative AGI timelines.
4. Limited exploration of risk-mitigating factors.
5. Possible conflation of different types of AI risks.
6. Insufficient use of rigorous probabilistic reasoning.
7. Limited consideration of potential positive AI scenarios.
8. Potential recency bias in weighing expert opinions.
9. Absence of sensitivity analysis for different assumptions.
10. Limited discussion of specific catastrophic scenarios.

The article concludes by suggesting improvements to the analysis, including expanding the range of queries, using more diverse perspectives, and implementing more rigorous probabilistic methods.[105]

---

[104] Driscoll, S. 'Is a "Rogue AI" Catastrophe Coming? Stanford's Mark Lemley on California's AI Safeguards Bill' SLS 25 September 2024 available at https://law.stanford.edu/2024/09/25/is-a-rogue-ai-catastrophe-coming-stanfords-mark-lemley-on-californias-ai-safeguards-bi/ (accessed 26/09/24)

[105] Golan, Y. 'Assessing AI Risks: AI Bets 50% on AI-induced Catastrophe Within 10 Years' Medium 25 September 2024 available at https://medium.com/@yanivg/assessing-ai-risks-ai-bets-50-on-catastrophe-within-10-years-d1ebd872e99b (accessed 26/09/2024)

**25 September 2024.** The Financial Times publishes an opinion piece by Anja Manuel, executive director of the Aspen Strategy Group, advocating for mandatory safety testing of large AI models.[106] Manuel draws parallels with pharmaceutical industry regulations, arguing that AI technologies like GPT-4 offer significant benefits but also pose severe national security risks.

The article highlights AI's potential to aid in creating biological or chemical weapons, emphasizing the need for testing to focus on "tangible, physical harms." Manuel cites recent incidents of state-sponsored hackers using OpenAI's technology for cyber attacks and references NATO's concerns about AI-powered terrorist attacks.

Manuel proposes a regulatory approach similar to drug safety laws, suggesting rapid testing within weeks to avoid hindering innovation. She recommends ongoing monitoring and reporting of model misuse by AI companies, balanced with "safe harbour" provisions to shield cooperating companies from some legal liabilities.

The piece positions both state and non-state actors as potential sources of AI-related threats, suggesting that AI could lower barriers for developing and using chemical, biological, and cyber weapons. It reflects growing concerns about AI's role in exacerbating CBW risks and calls for proactive governance measures to mitigate these threats while fostering responsible AI development.

**25 September 2024** - Frontiers in Bioengineering and Biotechnology publishes a study titled "Bridging biosafety and biosecurity gaps: DURC and ePPP policy insights from U.S. institutions" by Gillum et al.[107] The study provides empirical data on knowledge and practices related to dual use research of concern (DURC) and enhanced potential pandemic pathogens (ePPP) research across various U.S. sectors. It aims to improve oversight and inform policy development.

Key points:

- Government organizations were more likely to conduct DURC compared to other sectors
- Institutions with larger biosafety/biosecurity teams reported greater research activity and more effective non-compliance reporting mechanisms
- Public institutions were more likely to review experiments beyond the scope of the U.S. DURC Policy compared to private for-profit institutions
- Perceived financial support and challenges in policy implementation varied significantly across sectors

The study frames DURC and ePPP research as presenting unique biosafety and biosecurity hazards compared to less risky biological research. It positions both state actors (government institutions) and non-state actors (private companies, academic institutions) as potential sources of risk, while also highlighting their roles in risk mitigation.

The research emphasizes the importance of human resources, particularly the size of biosafety/biosecurity teams, in managing risks effectively. Financial resources are also highlighted as a key factor, with variations in perceived support across different types of institutions. These insights may be valuable when considering how AI applications interact with research ethics, processes, and outcomes, and the development of effective oversights. The study contextualizes its findings within broader discussions about balancing scientific innovation with biosafety and biosecurity concerns. It reflects growing awareness of the need for

[106] Manuel, A. 'It's time for limited, mandatory testing for AI' Financial Times 25 September 2024 available at https://www.ft.com/content/8b190ef1-32d8-4196-9643-3396a44d3bcd (accessed 30/09/2024)
[107] Gillum, D. R. et al. 'Bridging biosafety and biosecurity gaps: DURC and ePPP policy insights from U.S. institutions' (*Frontiers in Bioengineering and Biotechnology*, 2024) 12

robust oversight mechanisms and the challenges of implementing consistent practices across diverse institutional settings.

The authors implicitly assume that DURC and ePPP research will continue, focusing on how to manage risks rather than whether such research should be conducted at all. The study also reveals potential gaps in oversight, particularly in the private sector, where ePPP research may be occurring without adequate review. This study provides valuable empirical data on the current state of DURC and ePPP oversight in the U.S., highlighting critical areas for policy improvement and resource allocation to enhance biosafety and biosecurity practices across different types of institutions.

**27 September 2024.** The UK Ministry of Defence releases "Global Strategic Trends: Out to 2055" (GST 7), a comprehensive report examining potential future scenarios and global challenges.[108] While the report does not extensively discuss AI in relation to chemical and biological weapons (CBW), it does address these topics separately and in related contexts:

1. Weapons of Mass Effect: The report mentions "weapons of mass effect" as a potential future threat, which could include advanced chemical or biological weapons. It states, "An expansion in the number of nuclear-armed states fielding more powerful weapons, combined with new weapons of mass effect, could create new challenges." This suggests a concern about the proliferation and potential evolution of CBW, although it does not explicitly link this to AI.

2. AI as a Transformative Technology: AI is framed as a key driver of change with wide-ranging implications. The report notes, "Artificial intelligence could bring significant benefits in a wide range of socio-economic areas, regulatory frameworks may struggle to keep pace, presenting a growing risk to individuals and societies." In the context of security, AI is seen as potentially transformative: "As a result of technology advances, a range of fields, including transport and logistics, manufacturing, health care, food and energy production and communication, could all look significantly different by 2055."

3. Future Conflict and Technology: The report suggests that technological advances, which would include AI, could significantly alter the nature of combat power: "Economic challenges, demographic changes, green energy transitions and technology advances could see states and other actors pursuing widely different forms of combat power in the future, although mass and conventional means of power projection will remain important."

4. Data and Power: The report emphasizes the growing importance of data, which is closely tied to AI capabilities: "As the volume of global data generation grows and storage and processing become more efficient, data is likely to be increasingly essential for government and business decision-making. As a result, access to data is likely to be a key component of global power for both state and non-state actors."

5. Non-State Actors: The report highlights the increasing role of non-state actors in the security landscape, which could have implications for the development and use of advanced technologies including AI and potentially CBW: "An increasing range of security actors, including non-state, is likely to lead to a more congested and complex landscape out to 2055."

6. Regulatory Challenges: The report acknowledges the difficulty in regulating rapidly advancing technologies like AI: "regulatory frameworks may struggle to keep pace, presenting a growing risk to

[108] UK Ministery of Defence 'Global Strategic Trends out to 2055' HMG 27 September 2024 available at https://www.gov.uk/government/publications/global-strategic-trends-out-to-2055 (accessed 30/09/2024)

individuals and societies." This could have implications for controlling the development of AI-enhanced CBW.

The document frames these issues within a broader context of global power competition, technological advancement, and evolving security challenges. It suggests that both state and non-state actors could be involved in developing and potentially using advanced weapons technologies, including AI and possibly CBW.

While the report does not explicitly link AI and CBW, it does present them as part of a complex future security landscape where technological advances, including AI, could significantly alter the nature of threats and conflicts.

**27 September 2024.** The Center for Strategic and International Studies (CSIS) publishes a commentary by Zelie Petit titled "The Strategic Imperative of Biotechnology: Implications for U.S. National Security."[109] The piece extensively discusses the intersection of AI and biotechnology, including their potential applications in chemical and biological weapons development.

The commentary highlights how AI-powered biotechnology can facilitate the creation of novel biological warfare agents. It references a 2021 experiment where an AI drug discovery platform generated over 40,000 potentially toxic molecules, including known chemical warfare agents, in less than six hours. Petit notes:

"AI drug discovery platforms, which are often trained on open-source biological data, existing molecular structures, and associated proteins, could generate new molecular structures, and facilitate genomic targeting."

The article also discusses the potential for gene-editing technologies like CRISPR to create "precision genomic targeting biological warfare agents." It mentions concerns about China's collection and analysis of genomic data, which could potentially be used to develop targeted pathogens.

Petit frames these developments within the context of U.S.-China strategic competition, noting differing approaches to biotechnology development and regulation. The commentary emphasizes the dual-use nature of these technologies, highlighting their potential for both beneficial applications and weapons development.

This piece reflects growing concerns about AI's role in lowering barriers for the development of chemical and biological weapons, while also acknowledging its potential benefits in defense and other sectors. It underscores the complex challenges faced by policymakers in regulating and harnessing these emerging technologies.

**29 September 2024.** California Governor Gavin Newsom vetoes SB 1047, an artificial intelligence safety bill that would have established requirements for developers of advanced AI models to create protocols aimed at preventing catastrophes. The bill, introduced by Sen. Scott Wiener (D-San Francisco), faced fierce debate in Silicon Valley, with support from prominent AI researchers and opposition from major tech companies like Meta and OpenAI.[110]

The press release from Senator Wiener's office explicitly mentions the risk of AI being used to "develop chemical, nuclear or biological weapons" as one of the key threats the bill sought to address.[111] This framing

---

[109] Petit, Z. 'The Strategic Imperative of Biotechnology: Implications for U.S. National Security' CSIS 27 September 2024 available at https://www.csis.org/blogs/strategic-technologies-blog/strategic-imperative-biotechnology-implications-us-national (accessed 01/10/2024)

[110] Lee, W. 'Gov. Gavin Newsom vetoes AI safety bill opposed by Silicon Valley' Los Angeles Times 29 September 2024 available at https://www.latimes.com/entertainment-arts/business/story/2024-09-29/gov-gavin-newsom-vetoes-ai-safety-bill-scott-wiener-sb1047 (accessed 30/09/2024)

[111] 'Senator Wiener Responds to Governor Newsom Vetoing Landmark AI Bill' Senator Scott Wiener Press Release 29 September 2024 available at https://sd11.senate.ca.gov/news/senator-wiener-responds-governor-newsom-vetoing-landmark-ai-bill (accessed 30/09/2024)

positions AI as a potential enabler for the creation of weapons of mass destruction, including chemical and biological weapons. The statement from Geoffrey Hinton, former AI lead at Google, emphasizes the rapid and unpredicted progress of AI, implying that its potential for misuse, including in weapons development, may have outpaced regulatory efforts.

Nathan Calvin from the Center for AI Safety Action Fund warns of "catastrophic threats to society from AI," which could be interpreted to include the development of chemical and biological weapons, though this is not explicitly stated. The press release notes that 73% of AI researchers expressed "substantial" or "extreme" concern about AI falling into the hands of dangerous groups. While not directly linked to CBW, this concern could encompass fears about AI-assisted weapons development.

The veto of SB 1047 is presented as leaving a regulatory gap in addressing AI risks, potentially including those related to chemical and biological weapons. The press release argues that voluntary commitments from industry are insufficient to address these risks. This event highlights growing concerns about AI's potential role in lowering barriers for the development of chemical and biological weapons, and reflects the challenges in implementing regulatory measures to mitigate these risks. The framing suggests that both state and non-state actors could potentially exploit AI for weapons development, emphasizing the need for proactive governance measures.

## OCTOBER 2024

**2 October 2024** - Global Biodefense publishes an article titled "Tackling Misinformation and Distrust is Key to Improving Public Health Communication for the Next Pandemic" by Shauna Hurley and Rebecca Ryan, that discusses the importance of effective public health communication during pandemics, focusing on lessons learned from COVID-19 and other disease outbreaks.[112] It emphasizes the critical role of public trust and the need to counter misinformation, and that clear, consistent messaging from authorities is crucial for effective public health communication. Looking back, misinformation, especially on social media, posed a significant challenge during the COVID-19 pandemic and the article calls for coordinated international responses to counter misinformation and improve scientific literacy.

The article frames public health communication as a critical tool in managing biological threats, highlighting both its potential benefits when done effectively and the risks when it fails. While not explicitly discussing chemical or biological weapons, the emphasis on countering misinformation and building public trust has implications for how societies might respond to deliberate biological threats.

The piece positions government authorities and health institutions as key actors in shaping public responses to health crises. It also acknowledges the role of social media platforms and the general public in spreading or countering misinformation. The main resources discussed are informational - accurate health data, clear messaging, and efforts to improve public scientific literacy.

While there is no direct discussion of how these communication challenges might apply to deliberate biological attacks rather than natural outbreaks, the article does recognise AI has a role to play in information generation and dissemination.

[112] Hurley, S. and Ryan, R. 'Tackling misinformation and distrust is key to improving public health communication for the next pandemic' Global Biodefense 2 October 2024 available at https://globalbiodefense.com/2024/10/02/tackling-misinformation-and-distrust-is-key-to-improving-public-health-communication-for-the-next-pandemic/ (accessed 10/10/2024)

**4 October 2024** - The United Kingdom makes a voluntary contribution of £650,000 to the Organisation for the Prohibition of Chemical Weapons (OPCW), with a portion specifically allocated to AI-related projects.[113] In particular, £75,000 is dedicated to the OPCW AI Research Challenge project, which aims to identify ways AI can improve the efficiency and effectiveness of the OPCW's work. £50,000 supports the OPCW's Global Conference on AI in Advancing Implementation of the Chemical Weapons Convention, to be held in October 2024 in Morocco. It is noted that this conference will explore AI's role in chemical security, inform policy, and strengthen international collaboration. These AI initiatives are framed as part of efforts to address emerging threats and improve the OPCW's verification and inspection capabilities.

The contribution also includes funds for assistance and protection activities related to Ukraine (£200,000) and support for OPCW work in Syria (£200,000), including the Declaration Assessment Team, Fact-Finding Mission, and Investigation and Identification Team.

**5 October 2024** - Global Biodefense publishes an article titled "Safety Considerations for Chemical and Biological AI Models" discussing a Request for Information (RFI) issued by the U.S. Artificial Intelligence Safety Institute (AISI).[114] The AISI is seeking input on practices and methodologies for the responsible development and use of chemical and biological (chem-bio) AI models. The article highlights the dual-use nature of these technologies and the need to address potential risks.

Some key points include:

- Chem-bio AI models can accelerate progress in areas like drug discovery and medical countermeasures, but also pose risks if misused
- Examples of relevant models include protein design tools, genome assembly tools, and autonomous experimental platforms
- Concerns raised about AI potentially aiding in the design of more virulent pathogens or biological agents that can evade biosecurity measures
- The RFI seeks input on evaluation methodologies, risk assessment, and ways to strengthen existing biodefense measures

The framing emphasizes both the beneficial and potentially harmful applications of chem-bio AI models. It positions the issue as a dual-use challenge requiring proactive governance. The primary actors discussed are U.S. government agencies and the scientific community, with an emphasis on the need for collaboration. The resources highlighted are primarily technological and knowledge-based, focusing on AI capabilities and evaluation methodologies.

The article contextualizes the discussion within broader trends of rapid AI advancement and increasing convergence between AI and biotechnology. It implicitly assumes that AI development in this domain is inevitable, focusing the discussion on risk mitigation rather than prohibition. The emphasis on "responsible development" suggests a belief that proper governance can allow for innovation while managing risks.

The RFI and resulting discussion reflect growing awareness of AI's potential impact on chemical and biological security issues among policymakers and researchers. By soliciting broad input, it signals recognition of the complex, multidisciplinary nature of the challenge. This initiative represents an early attempt by the U.S.

---

[113] OPCW 'United Kingdom strengthens OPCW's global mission with £650,000 voluntary contribution to key activities' News 4 October 2024 available at https://www.opcw.org/media-centre/news/2024/10/united-kingdom-strengthens-opcws-global-mission-ps650000-voluntary (accessed 10/10/2024)

[114] 'Safety considerations for chemical and biological AI models' Global Biodefense 5 October 2024 available at https://globalbiodefense.com/2024/10/05/safety-considerations-for-chemical-and-biological-ai-models/ (accessed 10/10/2024)

government to proactively address the intersection of AI and chem-bio security risks through collaborative policy development.

**8 October 2024** - The European Union delivers a statement at the UN General Assembly First Committee's 79th session.[115]

The EU statement, delivered by Ambassador Hedda Samson, addresses a range of global security concerns, including specific mentions of chemical and biological weapons. Key points include:

- The EU expresses strong concern about reports of Russia allegedly using riot control agents as a method of warfare in Ukraine, which is prohibited under the Chemical Weapons Convention. They also note concerns about the alleged use of chloropicrin, a choking agent.
- The statement reaffirms the EU's commitment to strengthening multilateral instruments against chemical and biological weapons, emphasizing the need to uphold international prohibitions and ensure accountability for their use.
- The EU highlights the importance of the UN Secretary General's Mechanism in addressing chemical and biological weapons threats.
- Regarding emerging technologies, the statement mentions the need to address potential risks posed by lethal autonomous weapons systems (LAWS) and the responsible military use of artificial intelligence.

The EU frames chemical and biological weapons issues within a broader context of global security challenges, including Russia's war in Ukraine, conflicts in the Middle East, and the erosion of international norms. The statement positions the EU as a defender of international law and multilateral approaches to disarmament and non-proliferation.

While not directly addressing AI's potential role in chemical or biological weapons development, the EU's mention of AI in military contexts suggests growing awareness of its broader security implications.

**8 – 11 October 2024.** The Executive Council holds its 107th Session at OPCW headquarters in The Hague. For the first time, a regular item focusing on emerging technologies has been added to the agenda, as Agenda Item 10. The Annotated Provisional Agenda (EC-107/INF.1/Rev.1) notes, under this item: "The Secretariat circulated a Note by the Secretariat entitled "Artificial Intelligence and the OPCW: A Meeting with Experts" (S/2289/2024, dated 23 May 2024). The Council is requested to consider the matter." For references made by State Parties to AI during this session, please see References to AI in Sessions of the OPCW's PMOs.xlsx

**11 October 2024.** Ronith Lahoti publishes *The Socio-Ethical Dynamics of Artificial Intelligence in Healthcare* in *Significances of Bioengineering and Biosciences*. This article explores both the promise and challenges of AI in healthcare, focusing on socio-ethical considerations and the potential dual-use implications of AI in fields like clinical diagnostics, drug discovery, and protein engineering. Lahoti emphasizes the transformative role of AI technologies, including large language models (LLMs) and emerging quantum AI, in tasks such as medical diagnostics and personalized treatment planning. However, the author raises concerns about privacy, data governance, and the potential misuse of AI-driven protein engineering for non-peaceful purposes.

[115] 'EU statement – UN General Assembly 1st Committee: General Statement' EU 8 October 2024 available at https://www.eeas.europa.eu/delegations/un-new-york/eu-statement-%E2%80%93-un-general-assembly-1st-committee-general-statement_en (accessed 10/10/2024)

The article discusses the European Union's Artificial Intelligence Act as a regulatory framework aimed at managing AI risks, noting its classification of AI applications into risk categories, from minimal to unacceptable. Lahoti argues that while frameworks like the AI Act represent progress, they must evolve to address rapid technological advancements and dual-use risks. In particular, the article cautions that AI's capacity to autonomously process vast biomedical data may lower barriers to developing engineered biological agents, necessitating strict regulatory oversight.

To mitigate these risks, Lahoti calls for a collaborative approach involving policymakers, researchers, and healthcare practitioners. The recommendations include implementing transparency measures, fostering international cooperation, and ensuring AI applications align with ethical standards to avoid misuse while maximizing public health benefits. By drawing on the EU's regulatory approach and emphasizing AI's potential for both positive and negative impacts, the article advocates for a balanced, ethically guided development path for AI in healthcare.[116]

**17 October 2024.** The Nonproliferation Review publishes a comprehensive analysis by Stendall, Martin and Sandbrink examining how large language models could potentially lower technical barriers for chemical weapons development and use. The article presents a detailed assessment of AI's dual-use implications for chemical security, focusing particularly on how LLMs might enable actors to better reach "competency thresholds" across multiple domains of chemical weapon development.

The authors identify several key pathways through which LLMs could aid malicious actors:

- Enhanced ability to identify and understand chemical agents and their properties through "jailbreaking" techniques that bypass AI safety measures
- Improved capabilities for acquiring necessary materials through AI-enabled deception and creation of front companies
- Automated recruitment of skilled personnel through purpose-built chatbots
- Generation of technical guidance for chemical synthesis and dispersal methods
- Use of AI-powered social media bots to spread disinformation during attacks

The analysis frames these developments primarily through the lens of non-state actors and terrorist groups, positioning LLMs as potential "democratizers" of chemical weapons capabilities. The authors cite historical examples, including Aum Shinrikyo's 1995 Tokyo subway attack and chemical weapons use in Syria, to contextualize how reduced technical barriers could enable similar incidents.

The article identifies significant limitations and countermeasures, including: persistent financial barriers, requirements for specialized facilities and equipment, and potential uses of LLMs for chemical defense and counterterrorism. Three key policy recommendations are offered: enhanced security training for scientists, pre-release evaluations of LLMs for chemical security risks, and stricter regulation of chemical-focused AI tools.

This detailed examination highlights growing recognition that AI could fundamentally alter longstanding technical and knowledge barriers in chemical weapons development, while also suggesting pathways for risk mitigation through policy interventions.[117]

---

[116] Lahoti, R. 'The Socio-Ethical Dynamics of Artificial Intelligence in Healthcare' (*Significances of Bioengineering and Biosciences*, 2024) available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4982593 (accessed 11/11/2024)
[117] Stendall, R., Martin, F. and Sandbrink, J. 'How might large language models aid actors in reaching the competency threshold required to carry out a chemical attack?' Comment (*The Nonproliferation Review*, 2024) available at https://www.tandfonline.com/doi/full/10.1080/10736700.2024.2399308

**21 October 2024.** GEN Biotechnology publishes a perspective piece by Sean Ekins reflecting on a National Academies of Sciences, Engineering, and Medicine (NASEM) meeting examining the intersection of artificial intelligence and biosecurity. The article provides insights from key presentations at the August 2024 NASEM meeting, which brought together scientists from industry and academia to assess how AI could increase biosecurity risks, particularly around pandemic threats, while also exploring its potential benefits for reducing such risks.

Key points:

- Access to AI technology for molecule or protein design is available to "thousands of companies and scientists," with open-source software for generative molecule design accessible to "anyone" with minimal computational resources
- Google DeepMind's AlphaFold 3 demonstrates enhanced capabilities in protein structure prediction, though limitations remain around predicting viral pathogenicity
- Multiple speakers highlighted the implementation of safety measures in AI tools, including "filter agents" in CRISPR-GPT and proposals for "unlearning hazardous knowledge"
- The article notes that "biosecurity aspects came second to the science" in most presentations, with commercial interests potentially overwhelming security concerns

The discussion frames AI-CBW risks primarily through the lens of dual-use technology and democratized access. State actors receive limited attention, with focus instead on potential misuse by individuals or small organizations with relevant scientific knowledge. The resource landscape described spans from high-end commercial AI development (Google DeepMind, Microsoft) to readily accessible open-source tools, with human expertise presented as both an enabling factor and potential limiting barrier.

The article positions this discussion within broader debates about AI governance and biosecurity, referencing President Biden's Executive Order 14110 and ongoing policy discussions. However, it notably omits detailed consideration of international dimensions or existing CBW control regimes, instead focusing primarily on U.S. domestic policy responses and technical solutions. The author expresses skepticism about the likelihood of meaningful regulation emerging from these discussions, noting that "it is unlikely to lead to any regulations that will impact AI use and biosecurity unless the panel members in closed sessions go beyond what was presented."

This milestone meeting and subsequent analysis highlight growing recognition of AI's dual-use potential in biological and chemical contexts, while also revealing significant gaps between technical capabilities, security concerns, and governance responses.[118]

**21 October 2024.** While this document does not directly address chemical or biological weapons, it contains significant discussion of AI integration with IoT technologies and associated security concerns. The report focuses on recommendations for securing and advancing IoT infrastructure while acknowledging emerging risks from the convergence of AI and IoT systems. The document frames AI primarily as both an enabler and potential risk factor for IoT systems. Key quotes include:

"AI requires and incorporates the use of data from various sources to build and train models, as well as make decisions and act upon those decisions"

[118] Ekins, S. 'Biosecurity and Artificial Intelligence in the Life Sciences' (*GEN Biotechnology*, 2024) 3(5) available at https://www.liebertpub.com/doi/10.1089/genbio.2024.0051 (accessed 31/10/2024)

"The convergence of AI with IoT (AIoT) is poised to drive transformation across wide sectors of the economy, but its development and use must be managed to foster the proper outcomes and minimize unintended consequences."

The report identifies several critical resources and concerns:

- Data Infrastructure: Notes that by 2025, there will be 55.9 billion IoT devices generating 79.4 zettabytes of data
- Security Vulnerabilities: Highlights risks from "AI-based cyberattacks" and the need for "quantum-safe cryptographic methods"
- Technical Capabilities: Discusses edge computing and AI-capable chips as crucial infrastructure

The document places these developments within broader geopolitical contexts, particularly noting:

- Concerns about Chinese dominance in IoT modules (64% global market share)
- Supply chain vulnerabilities and national security implications
- Need for international cooperation on standards and security

Significant omissions/assumptions:

- Limited discussion of military applications of AIoT systems
- No explicit consideration of dual-use potential for AIoT technologies
- Minimal discussion of biological/chemical sensing applications, though environmental and health monitoring are mentioned

The report's emphasis on securing AIoT infrastructure and maintaining technological competitiveness suggests growing recognition of these systems' critical importance to national security, though it notably avoids direct discussion of military or weapons applications.[119]

**21 October 2024.** A Congressional Research Service report examining approaches to federal data standardization highlights the growing complexity of managing data standards across government agencies, with notable implications for CBW-related information sharing and security. The report arrives amid increasing concerns about data governance in sensitive domains, including those affecting national security and weapons proliferation.

The analysis unpacks several critical challenges facing federal data management. Central to these is the tension between program-specific standards and government-wide requirements - an issue particularly relevant for agencies dealing with CBW-related information. While the report doesn't explicitly focus on CBW matters, its discussion of data standards in sensitive contexts like homeland security offers important insights for how agencies might better coordinate CBW-related data sharing and analysis.

Of particular note is the Department of Homeland Security's Data Framework Act implementation, which the report uses to illustrate challenges in integrating existing datasets and systems across agency components. This example resonates strongly with ongoing efforts to coordinate CBW-related intelligence and monitoring across different agencies and jurisdictions.

The report reveals an evolving landscape of federal data governance, where Chief Data Officers increasingly play crucial roles in managing sensitive information. Their responsibilities - from developing data standards to facilitating common languages for data stewards - have direct bearing on how agencies might better coordinate CBW-related monitoring and response capabilities.

---

[119] NIST 'Internet of Things (IOT) Advisory Board (IoTAB) Report' 21 October 2024, available at https://www.nist.gov/system/files/documents/2024/10/21/The%20IoT%20of%20Things%20Oct%202024%20508%20FINAL_1.pdf (accessed 31/10/24)

A striking omission in the analysis is direct discussion of how data standardization affects the handling of weapons-related information across agencies. The report's focus on more general administrative and financial data leaves open questions about how its recommendations might apply to more sensitive security domains.

This analytical framework arrives at a critical moment when AI and other emerging technologies are generating vast new datasets requiring standardization and governance, with profound implications for how federal agencies monitor and respond to CBW-related threats.[120]

**22-24 October 2024.** The Kingdom of Morocco hosts the OPCW's Global AI Conference in Rabat. A report will be forthcoming.[121]

**22 October 2024.** The U.S. Government Accountability Office (GAO) releases a comprehensive report examining commercial development and deployment practices of generative AI technologies. The report, responding to congressional requests, highlights key security concerns around AI development, including specific references to risks related to chemical, biological, radiological, and nuclear (CBRN) weapons.

Key points:

- Commercial developers conduct "red teaming" to test AI systems specifically for "chemical, biological, radiological, and nuclear risks," following requirements in the October 2023 White House Executive Order
- The report identifies "prompt injection" attacks as a significant concern, noting that users can potentially circumvent AI safety measures by reframing prohibited requests (like weapon-making instructions) in ways that bypass security controls
- Attackers may not need advanced technical knowledge to exploit AI systems, with the report noting that "these attacks do not require advanced programming knowledge or technical skills"
- Despite implementation of safety measures, developers acknowledge their models remain susceptible to generating harmful outputs that could have "significant consequences"

The report frames AI security risks primarily through the lens of commercial development and cybersecurity, positioning both state and non-state actors as potential threats. Notable resources discussed include technical safeguards (red teaming, monitoring systems), human resources (multi-disciplinary teams), and data resources (training datasets). The report emphasizes how readily accessible these technologies are becoming, with one developer reporting "more than 200 million weekly active users."

The GAO analysis sits within broader U.S. government efforts to understand and regulate AI technologies, building on the White House's October 2023 Executive Order while focusing specifically on commercial development practices. Notably, while the report acknowledges CBRN risks, it does not deeply examine the specific mechanisms by which AI might enable such weapons development, instead treating them as one category among many potential security concerns. The report's framing suggests an implicit assumption that technical and procedural safeguards, if properly implemented, can adequately address these risks.

The GAO's findings highlight growing recognition within oversight bodies of AI's potential to enable CBRN weapons development, while also revealing gaps between commercial safeguards and emerging security challenges.[122]

---

[120] Congressional Research Service 'Standardising Federal Data: Categorizing Approaches' 21 October 2024 available at https://crsreports.congress.gov/product/pdf/IF/IF12787 (accessed 31/10/24)
[121] OPCW 'Global Conference on AI in CWC implementation' https://www.opcw.org/media-centre/featured-topics/aiconference (accessed 31/10/2024)
[122] US GAO 'Artificial Intelligence: Generative AI Training, Development, and Deployment Considerations' GAO-25-107651, 22 October 2024, available at https://www.gao.gov/products/gao-25-107651 (accessed 31/10/2024)

**22 October 2024**. Gartner releases its annual "Top Strategic Technology Trends" report for 2025, highlighting AI governance and disinformation security among key emerging challenges. The report specifically notes AI's potential role in chemical, biological, radiological and nuclear risks, positioning these concerns within broader technological developments that will shape enterprise security.

In discussing AI security challenges, Gartner emphasizes several converging threats. The report warns of AI's increasing autonomy through "agentic AI" systems that can independently make decisions and take actions to achieve goals. By 2028, the analysts predict that at least 15% of day-to-day work decisions will be made autonomously through such systems - marking a dramatic shift from zero percent in 2024. This growing AI autonomy raises particular concerns around weapons development and security risks.

The report places special emphasis on disinformation security, describing it as a "digital arms race" where AI-enhanced phishing, hacktivism and social engineering pose escalating threats. Gartner projects that by 2028, half of enterprises will need to adopt specific products and services to combat AI-enabled disinformation - a dramatic increase from less than 5% in 2024. This trend has significant implications for CBW-related risks, particularly around the spread of technical knowledge and potential exploitation of AI systems.

Gartner's analysis frames these developments within a broader context of evolving AI governance needs. The report advocates for comprehensive governance platforms that can ensure AI systems remain "reliable, transparent, fair and accountable while also meeting safety and ethical standards." This governance emphasis reflects growing awareness that AI applications in sensitive domains, including potential CBW applications, require robust oversight mechanisms.

The report also reveals an implicit tension between AI's commercial benefits and its security risks. While acknowledging serious concerns around AI misuse, including CBRN risks, the analysis primarily focuses on enterprise applications and competitive advantages. This framing suggests an ongoing challenge in balancing innovation with security considerations - particularly relevant for dual-use AI technologies with potential CBW applications.

The report demonstrates the mainstreaming of AI security concerns within enterprise technology planning while highlighting significant gaps in current governance approaches.[123]

## NOVEMBER 2024

**6 November 2024.** The UK government launches a new platform to help businesses assess and mitigate AI risks. The platform provides guidance and resources for conducting impact assessments, evaluating new AI technologies, and reviewing machine learning algorithms for bias. Science and Technology Secretary Peter Kyle frames this initiative as part of the UK's ambition to become "a true hub of AI assurance expertise," building on the work of the UK AI Safety Institute established under the previous Conservative government.

The initiative positions AI safety and assurance technology as a strategic niche for the UK, with the government estimating the domestic market could grow to £6.5bn by 2035. The announcement includes a new partnership with Singapore for cooperation between both countries' safety institutes on research and standards development. The platform includes a self-assessment tool specifically designed for small businesses to verify safe AI system use.

The article makes no specific mention of chemical or biological weapons, or military applications of AI, although its focus on commercial and economic aspects of AI safety and regulation speaks to infrastructures that may be relevant to questions of acquisition and chemical/bio safety and security.[124]

---

[123] Gartner 'Top 10 Stategic Technology Trends for 2025' 22 October 2024 available at https://www.gartner.com/en/articles/top-technology-trends-2025 (accessed 31/10/2024)

[124] Gross, A. 'UK government launches new AI safety platform for businesses' Financial Times, 6 November 2024, available at https://www.ft.com/content/8a54932d-d9a9-4a69-969d-89d8b2de149f (accessed 6/11/2024)

**6 November 2024.** Shravishtha Ajaykumar publishes an occasional paper titled *Strengthening CBRN Security in India: Domestic Strategies and Global Collaborations* through the Observer Research Foundation. The paper addresses India's strategic need to enhance CBRN (Chemical, Biological, Radiological, and Nuclear) security, given regional security challenges and advancing technologies. Ajaykumar highlights the growing role of artificial intelligence (AI) and machine learning (ML) in bolstering CBRN defenses. AI applications such as predictive modeling, biosurveillance, and real-time data analysis are emphasized for their capacity to rapidly identify, monitor, and respond to potential biological and chemical threats. Additionally, the paper underscores the importance of AI-driven early warning systems and automated response strategies to improve detection accuracy and operational speed in crisis situations. Ajaykumar recommends that India integrate AI across its CBRN strategy while addressing potential cybersecurity gaps. Strengthening international alliances, particularly within the Quad, and fostering public-private partnerships are highlighted as critical for driving innovation and ensuring resilience in India's CBRN defense.[125]

**25 – 29 November 2024**. Held at OPCW headquarters in The Hague, the Netherlands, the 29th Session of the OPCW's Executive Council contained a number of references to AI by both the Director-General and States Parties, signalling the increasing attention the subject is receiving. A compilation of these remarks is available as a standalone file.

## DECEMBER 2024

**4 December 2024.** The James Martin Center for Nonproliferation Studies publishes *"Generative AI and WMD Nonproliferation: A Practical Primer for Policymakers and Diplomats"* by Natasha E. Bajema. The report examines the implications of generative artificial intelligence (AI) for weapons of mass destruction (WMD) nonproliferation, with a significant focus on chemical and biological weapons (CBW). It highlights the dual-use nature of generative AI, identifying risks and opportunities for the CBW domain:

- **Risks:** Generative AI systems could lower barriers for malicious actors in CBW development by facilitating the design of toxic chemicals or pathogens. Tools like AlphaFold, originally developed for medical research, might be repurposed to predict and optimize the properties of harmful biological agents.
- **Opportunities:** The same technologies could support nonproliferation efforts by enhancing monitoring and verification processes, analyzing complex datasets, and detecting patterns indicative of proliferation activities.

The report underscores the challenges of assessing AI's role in CBW risks due to the technology's novelty and complexity. It calls for:

1. **Benchmarking AI Capabilities:** Establishing metrics to evaluate AI's potential contributions to WMD proliferation risks.
2. **Regulatory Action:** Implementing oversight mechanisms, particularly for open-source AI models, to mitigate dual-use concerns.
3. **International Collaboration:** Aligning governance frameworks with global nonproliferation treaties, such as the Chemical Weapons Convention (CWC) and Biological Weapons Convention (BWC).

[125] Ajaykumar, S. 'Strengthening CBRN Security in India: Domestic Strategies and Global Collaborations' *Occasional Paper No. 452* (Observer Research Foundation, 2024)

This primer situates generative AI as both a potential disruptor and a tool for reinforcing CBW nonproliferation, emphasizing the need for proactive governance to balance risks and benefits.[126]

**5 December 2024.** The Conversation publishes an article titled *"Chatbots won't help anyone make weapons of mass destruction – but other AI systems just might,"* by David Heslop and Joel Keep.

The article distinguishes between the risks posed by general-purpose AI, such as large language models (LLMs), and specialized scientific AI systems in enabling chemical and biological weapons (CBW) development. Heslop and Keep argue that while chatbots like ChatGPT may help navigate public data, they lack the capability to meaningfully advance bioterrorism efforts. In contrast, AI tools designed for molecular research, such as AlphaFold, present dual-use risks by potentially aiding the synthesis of dangerous proteins or toxins.

The authors highlight regulatory efforts under the Biden administration's 2023 Executive Order on AI, which seeks to address the proliferation risks of CBRN+AI technologies and regulate tools like gene synthesis hardware. They caution against rolling back these nascent frameworks, stressing their importance for global biosecurity.[127]

**12 December 2024.** The Carnegie Endowment for International Peace publishes a report by Holden Karnofsky titled *"A Sketch of Potential Tripwire Capabilities for AI."* The report proposes the establishment of specific "tripwire capabilities" to identify and mitigate risks associated with advanced AI systems before they reach critical thresholds. Among the tripwires suggested is a framework to assess when AI models could substantially assist malicious actors in developing chemical and biological weapons (CBW).

Karnofsky highlights scenarios where AI could lower barriers to CBW development. He posits that AI could act as a "virtual expert," providing operational guidance that bridges technical gaps, such as chemical synthesis, weaponization, or dissemination methods. This potential, the report suggests, could democratize access to CBW capabilities, extending risks to non-state actors and individuals who lack traditional expertise.

Key recommendations include:

- **Defining Risk Thresholds:** Establishing metrics to evaluate when AI systems demonstrate competencies that pose a significant threat to CBW proliferation.
- **Mandatory Safety Testing:** Implementing pre-release evaluations of AI systems to determine whether they meet tripwire thresholds for CBW risks.
- **Global Governance Coordination:** Encouraging international collaboration to align tripwire standards with existing frameworks, such as the Biological and Chemical Weapons Conventions.

The report also emphasizes the need for transparency among AI developers, calling for external audits and red-teaming exercises to identify vulnerabilities in AI models before their deployment. Karnofsky's approach

[126] Bajema, N. 'Generative AI and WMD Nonproliferation: A Practical Primer for Policymakers and Diplomats' CNS Occasional Paper #63 December 2024 available at: https://nonproliferation.org/wp-content/uploads/2024/12/generative_ai_and_wmd_nonproliferation_12042024.pdf (accessed 20/01/2025)
[127] Heslop, D. and Keep, J. 'Chatbots won't help anyone make weapons of mass destruction – but other AI systems just might' The Conversation, 5 December 2024 available at: https://theconversation.com/chatbots-wont-help-anyone-make-weapons-of-mass-destruction-but-other-ai-systems-just-might-244514 (17/01/25)

reflects a precautionary stance, advocating for proactive measures to prevent the misuse of transformative AI technologies in CBW contexts.[128]

[128] Karnofsky, H. 'A Sketch of Potential Tripwire Capabilities for AI' Carnegie, December 2024. Available at https://carnegieendowment.org/research/2024/12/a-sketch-of-potential-tripwire-capabilities-for-ai?lang=en (accessed 17/01/2025