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

Draft, version 28 March 2024

Introduction

This chronology presents a detailed timeline of events and policy documentation of relevance to broader discussions around Artificial Intelligence in the Chemical and Biological Weapons Prohibition Regimes. This chronology is still 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 key parameters:

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

Entries satisfying the first two criteria were included; those which satisfy the third criteria but not the second were provisionally labelled [UI], or Uncertain of Inclusion. This enabled us to highlight entries which require some discussion as to how these may have relevance now and in the future. Additional labels, particularly where an entry is light on detail, include:

- [H] – Holding, we know the event happened, but need to find and add detail
- [Q] – Query, we have some detail on the event but there are some issues in the manner of its reporting, or broader questions about the event's implications.

Note on sources

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.

Archiving our source materials

All documents/files begin with date format YYYY-MMDD, and this date correlates the date of publication *and* the entry in the chronology. There may be cases where the document does not align with the event itself. For example a Conference may be held on 1 May, but the report which is our primary reference document is not published until 15 May. In these cases, the document is archived with the *date of the entry*, as the document informs our understanding of the event itself.

Then, { } designates the country, organisation or company the document originates from. When designating a country, use the ISO 3166-1 alpha-2 codes for each country (i.e. the internationally recognised 2 letter code). Codes can be found here. Designating the country is particularly important for documents that come from states, their diplomatic delegations, their government departments and agencies, and so forth; it is also important for newspapers as this speaks to local media consumption contexts.

Then, [ ] designates the specific subsidiary body, organ, department, news organisation, that is generating the document. The [ ] can be used utilised twice in succession, if relevant to further specify into which forum the document is being released into. This is particularly relevant when a country makes a statement at an international organisation in one of its subsidiary organs. It may also be useful if an additional categorisation is needed.

For example:

1. A statement by the United States of America at the OPCW for the 5th Review Conference
    2023-1405 {US} [OPCW] [RC5] Statement on improved verification techniques
2. A news article in Le Monde
    2023-0802 {FR} [LeMonde] Je suis AI
3. An article by Microsoft on their online magazine Pulse
    2023-0802 {Microsoft} [Pulse] AI in the workplace, what's new

CHRONOLOGY ENTRIES

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 titled "Emerging Technologies and the CWC: Autonomous Systems and Artificial Intelligence" [1]

[1] Presentation available at: https://www.opcw.org/sites/default/files/documents/Science_Technology/Diplomats_Programme/20151130-Autonomous_systems_and_AI-JForman-Print.pdf

[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. [1]

[1] https://www.defense.gov/News/News-Stories/Article/Article/1755942/dod-unveils-its-artificial-intelligence-strategy/

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. [1]

[1] The SAB'Report of the Scientific Advisory Board at its Twenty-Eighth Session 10-14 July 2019' SAB-28/1 14 July 2019

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" [Is the EU/EC doing this diplomacy work? Is the EC/EU making statements in other settings around AI and risks/security/CBW?]

[1] 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)), available at: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html

10-15 September 2021. The fourth Spiez Convergence conference is held online, with only rapporteurs and chairs gathering in person. The aim of the Spiez Convergence conference series is to continue discussions on the convergence of the sciences initiated by the OPCW's Scientific Advisory Board. AI has received attention, with the 2021 Conference Report's Executive Summary stating:

"Artificial Intelligence (AI) has become an important technology for the synthesis and utilisation of chemical molecules. The combination of improved algorithms, increased computational power as well

as open access to data is becoming a game changer and is making so far unknown molecules and entire "chemical spaces" accessible. A very promising field is the use of machine learning (ML). ML algorithms predict properties based on existing data to prioritise drugs for *in vitro* and *in vivo* testing. Curated data depositories are combined with trained algorithms to function as generative models, working like a medicinal chemist. Such a generative model could however also be employed to propose structures for toxic agents – an example based on the nerve agent VX was presented at the workshop. Using ML, the steps from molecular design to synthesis are becoming easier and they can be automated, with the downside, that such ML methods could be deployed to actively avoid control measures. The number of companies that are active in the field of AI is growing rapidly and so is the capital investment in this industry. The conference discussed an emerging risk in the AI industry due to a seeming absence of awareness about the misuse potential of AI as well as a lack of oversight." (p.8) [1]

[1] For links to all Convergence reports, see: https://www.spiezlab.admin.ch/en/home/meta/ref convergence.html

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 [1]

- "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" [2] [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?]

[1] 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

[2] 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

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 "the potential for dual-use" with such AI systems.[1] 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.[2][3]

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

[2] Much to discuss in AI ethics. Nat Mach Intell 4, 1055–1056 (2022). https://doi.org/10.1038/s42256-022-00598-x

[3] 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

16-17 June 2022. The OPCW's Scientific Advisory Board and the International Union of Pure and Applied Chemistry (IUPAC) jointly organise an international workshop on 'Artificial Intelligence Assisted Chemistry', held in the Hague, Netherlands. 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. [1] 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 specifically has recently published a journal article on responsible research and innovation of AI in the food industry.[2]

[1] 'International workshop on Artificial Intelligence Assisted Chemistry' Chemical Weekly 5 July 2022 available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjByuyC-r2AAxWHmKQKHTXgAKcQFnoECA0QAQ&url=https%3A%2F%2Fiupac.org%2Fwp-content%2Fuploads%2F2022%2F07%2FInternational_workshop_on_Artificial_Intelligence.pdf&usg=AOvVaw0IimTGWHpzzhUaN3VDxjxB&opi=89978449

[2] 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

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. [1] 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 within the military. These combined efforts aim to ensure more effective detection, response, and protection against CBRN threats while reducing unnecessary battlefield restrictions. [2]

[1] Pentagon Biological Defense Programs at "Pivot Point," https://www.nationaldefensemagazine.org/articles/2022/7/28/pentagon-biological-defense-programs-at-pivot-point

[2] Military Times. Big Changes Ahead for How Troops Battle Future Chemical, Biological

Threats. Retrieved from https://www.militarytimes.com/news/2022/08/02/big-changes-ahead-for-how-troops-battle-future-chemical-biological-threats/

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) [1]

[1] For links to all Convergence reports, see: https://www.spiezlab.admin.ch/en/home/meta/refconvergence.html

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 the population at large." [1]. 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.)." [2] Thus the handbook provides a good reference for the growing importance of understand AI as an ideational mediator through which challenges to the CBW Conventions can be mounted.

[1] UNICRI News Item 'New! UNICRI releases the handbook to combat CBRN disinformation' 13 January 2023 available at: https://unicri.it/News/Hanbook-to-combat-disinformation

[2] UNICRI 'Handbook to combat CBRN disinformation' 13 January 2023 available at: https://unicri.it/sites/default/files/2023-01/Handbook%20to%20combat%20CBRN%20disinformation.pdf

17 Feb 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." [1]

[1] Vigliarolo, Brandon. "International Military AI Summit Ends with 60-State Pledge." The Register® - Biting the Hand That Feeds IT, The Register, 18 Feb. 2023, www.theregister.com/2023/02/17/military_ai_summit/.

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 existing global arms control regimes, including the Nuclear Non-Proliferation Treaty and the Chemical and Biological Weapons Conventions. [1]

[1] Nato. "NATO Secretary General: We Stand at a Crossroads for International Arms Control." NATO, 18 Apr. 2023, https://www.nato.int/cps/en/natohq/news_213954.htm?selectedLocale=en
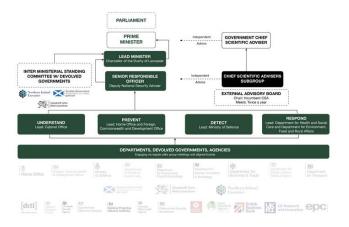
7 June 2023. The UK Government announces it will host "the first global summit on Artificial Intelligence" in autumn 2023, which "will consider the risks of AI, including frontier systems, and discuss how they can be mitigated through internationally coordinated action. It will also provide a platform 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 [...]." [1] Palantir has provided a range of AI-driven services to law enforcement agencies and militaries, including in the Russo-Ukrainian war. [2]

[1] UK HMG Press Release 'UK to host first global summit on Artificial Intelligence' 7 June 2023 available at: https://www.gov.uk/government/news/uk-to-host-first-global-summit-on-artificial-intelligence

[2] Ignatius, D. 'How the algorithm tipped the balance in Ukraine' The Washington Post 19 December 2022 available at: https://www.washingtonpost.com/opinions/2022/12/19/palantir-algorithm-data-ukraine-war/

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". [1] 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 possible threats.

*[1] The UK Biological Security Strategy 2023 is available at:
https://www.gov.uk/government/publications/uk-biological-security-strategy/uk-biological-security-strategy-html#executive-summary.*

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 controls at critical points are suggested

measures to address this risk, though challenges in implementing these safeguards exist. [1]

[1] Esvelt, K. (2023). Could Chatbots Help Devise the Next Pandemic Virus? Science. https://www.science.org/content/article/could-chatbots-help-devise-next-pandemic-virus

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 [1].

[1] United Nations (2023) Report of the Secretary-General on the Responsibility to Protect: Implementing the Responsibility to Protect: Accountability for Prevention. Available at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N23/189/56/PDF/N2318956.pdf?OpenElement

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. [1]

[1] https://www.telegraph.co.uk/global-health/science-and-disease/chatgpt-google-bard-ai-bioweapon-pandemic/

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 [1]. *The Strategy for Public Health Preparedness and Response to Artificial Intelligence Threats,* sponsored by Senators Markey and Budd [2]. *The Securing Gene Synthesis Act*, sponsored by Representative Eshoo and Senator Markey [3].

[1] H.R.4704 – 118th Congress (2023-2024): "Artificial Intelligence and Biosecurity Risks Assessment", H.R.4704, 118th Cong. (2023)

[2] S.2399 – 118th Congress (2023-2024): "Strategy for Public Health Preparedness and Response to Artificial Intelligence Threats", S.2399, 118th Cong. (2023)

[3] 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. [1]

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

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." [1]

[1] 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*

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" [1]. 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.

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. [1]

[1] https://www.reuters.com/technology/us-senators-express-bipartisan-alarm-about-ai-focusing-biological-attack-2023-07-25/

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. [1]

[1] https://documents-dds-ny.un.org/doc/UNDOC/GEN/G23/159/12/PDF/G2315912.pdf?OpenElement

9 August 2023. In Reedley, California, a small city in the Central Valley, the discovery of an unregistered medical lab has sparked rumors 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. [1]

[1] 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.

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. [1]

[1] https://www.unicc.org/news/2024/02/05/unicc-collaborates-with-unidir-and-vertic-to-develop-a-bwc-national-implementation-measures-database

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. [1]

[1] *https://documents-dds-ny.un.org/doc/UNDOC/GEN/G23/167/77/PDF/G2316777.pdf?OpenElement*

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.[1]

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

12 September 2023. 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.[1]

[1] "Executive Order on Advancing Biotechnology and Biomanufacturing Innovation for a Sustainable, Safe, and Secure American Bioeconomy." The White House, The United States Government, 12 Sept. 2022, 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/.

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. [1]

*[1]*https://www.theguardian.com/technology/2023/sep/25/ai-bioweapons-rishi-sunak-safety

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. [1]

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

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.[1][2] while other news organizations such as Sky News published headline titled "Mohamad al Bared guilty of terror offence after designing 'kamikaze' drone for ISIS".[3]The difference of headlines displays the difference of focus by each news organizations with some emphasizing the possibility of the created drone while other news sources focused more on what was made.

[1] https://www.bbc.co.uk/news/uk-england-coventry-warwickshire-66947311

[2] https://www.businessinsider.com/mohamad-al-bared-3d-print-drone-isis-chemical-weapon-2023-9?r=US&IR=T

[3] Duncan Gardham, security journalist. "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.

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. [1][2]

[1] "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.

[2] The Operational Risks of Ai in Large-Scale … - Rand Corporation, www.rand.org/pubs/research_reports/RRA2977-1.html. Accessed 18 Oct. 2023.

19 October 2023. The Testing Playbook for Biological Emergencies, authored collaboratively by subject matter experts at the Pandemic Center at the Brown University School of Public Health, Arizona State University College of Health Solutions, and Illumina Ventures, in partnership with the Association of Public Health Laboratories and the STAT Public Health Network at the Brown University School of Public Health, is a comprehensive guide designed to aid U.S. decision-makers at all levels of government in efficiently responding to infectious disease emergencies through diagnostic testing. It encompasses a wide range of testing methods, from individual diagnostics to group surveillance, at-home testing, and wastewater surveillance. Recognizing that testing plays a pivotal role in pandemic preparedness and response, the playbook offers a strategic approach to ensure equitable and rapid access to testing during outbreaks. The playbook is divided into six phases of a biological emergency, each with specific steps to ensure early access to accurate testing, scaling testing capabilities, and broad access to testing modalities. It underscores the importance of national coordination among various sectors to effectively contain disease spread and save lives. [1]

[1] "Testing Playbook for Biological Emergencies." Better Testing Now, www.bettertestingnow.org/testing-playbook-for-biological-emergencies. Accessed 2 Nov. 2023.

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.[1]

[1] "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.

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.[1]

[1] 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.

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, 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.[1]

[1] 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.

26 October 2023. In a BBC report on his recent address, Prime Minister Rishi Sunak raised concerns about the potential risks associated with artificial intelligence (AI), specifically its potential contribution to the development of chemical and biological weapons. In a worst-case scenario, Sunak warned of the possibility that society could lose control over AI, preventing it from being switched off. He emphasized the need to acknowledge and address the disputed harm potential of AI, cautioning against ignoring the risks by stating, "While the potential for harm is disputed, we must not put our heads in the sand over AI risks." In his effort to position the UK as a global leader in AI, Sunak acknowledged the technology's positive impact on job creation and economic growth. Despite recognizing its role in catalyzing economic growth and productivity, Sunak admitted that AI's development would have implications for the labor market, particularly in terms of job disruption. He stated, "The technology is already creating jobs. [Its] development will catalyze economic growth and productivity, though it will have an impact on the labor market." Referring to the capabilities and potential risks outlined in a government report on AI, Sunak highlighted the report's warning that AI could be exploited by terrorist groups to "spread fear and disruption on an even greater scale." While emphasizing the importance of addressing the risk of human extinction from AI as a "global priority," he balanced his message by stating, "This is not a risk that people need to be losing sleep over right now, and I don't want to be alarmist." Sunak further discussed the immediate threat posed by AI to jobs, citing examples of AI tools efficiently performing administrative tasks traditionally carried out by employees. He encouraged viewing AI as a "co-pilot" in day-to-day workplace activities rather than a complete replacement for human roles, asserting, "It is too simple to say artificial intelligence would take people's jobs." Reports, including the government's "Safety and Security Risks of Generative Artificial Intelligence to 2025" report, were mentioned, outlining potential threats such as enhancing terrorist capabilities, increasing fraud, child sexual abuse, cyberattacks, erosion of trust in information, and the use of deepfakes to influence societal debates. Rashik Parmar, the chief executive of the BCS, The Chartered Institute for IT, offered a perspective, stating, "AI won't grow up like The Terminator. If we take the proper steps, it will be a trusted co-pilot from our earliest school days to our retirement." Regarding AI regulation, Sunak emphasized a cautious approach, stating that the UK would not "rush to regulate" AI due to the difficulty of regulating something not fully understood. He advocated for proportionate regulation while encouraging innovation, asserting, "The UK's approach should be proportionate while also encouraging innovation." Despite challenges, Sunak aimed to position the UK as a global leader in AI safety. He acknowledged the difficulty of competing with major players like the US and China but emphasized the importance of UK leadership. Regarding the secrecy surrounding AI development, he expressed the need to persuade firms to stop "marking their own homework." In response to criticism, including concerns raised by former Prime Minister Liz Truss about inviting China to an AI safety summit, Sunak defended the decision, arguing that 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 in Buckinghamshire was highlighted, where world leaders, tech firms, scientists, and academics will gather to discuss the emerging technology. Professor Gina Neff, Director of the Minderoo Centre for Technology and Democracy at the University of Cambridge, criticized the focus of the summit, stating, "The concerns that most people care about are not on the table, from building digital skills to how we work with powerful AI tools. This brings its own risks for people, communities, and the planet." [1]

[1] James Gregory & Zoe Kleinman, technology editor. "Rishi Sunak Says AI Has Threats and Risks - but Outlines Its Potential." BBC News, BBC, 26 Oct. 2023, www.bbc.co.uk/news/uk-67225158.

29 October 2023. The UK Prime Minister has unveiled a new mission to accelerate the use of artificial intelligence(AI) in life sciences to address major health challenges. A £100 million government investment will be directed towards leveraging AI for

breakthroughs in treatments for previously incurable diseases. The AI Life Sciences Accelerator Mission capitalizes 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 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. [1]

[1] 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.

30 October 2023. NTI | bio is organizing the launch of their newest report, "The Convergence of Artificial Intelligence and the Life Sciences: Safeguarding Technology, Rethinking Governance, and Preventing Catastrophe." This comprehensive report, compiled through consultations with more than 30 experts spanning AI, biotechnology, bioscience research, and biosecurity, focuses on the pivotal intersection of AI and the life sciences. It not only addresses the biosecurity implications of this convergence but also outlines strategies to protect these technologies. The report offers a set of crucial recommendations for governance approaches, aiming to secure AI-bio capabilities while mitigating biological risks. This approach strikes a balance between ensuring safety and fostering scientific advancements. [1]

[1] "The Convergence of AI and the Life Sciences: New Report on Safeguarding Technology, Rethinking Governance, and Preventing Catastrophe - Ai Fringe." - AI Fringe, aifringe.org/events/the-convergence-of-ai-and-the-life-sciences-new-report-on-safeguarding-technology-rethinking-governance-and-preventing-catastrophe. Accessed 18 Oct. 2023.

30 October 2023. In the past seven weeks, the UK's Frontier AI Taskforce has achieved significant milestones, including tripling the research team's capacity with 150 years of AI research experience. Notable experts in safety protocols and governance, Jade Leung and Rumman Chowdhury, have been recruited, further strengthening the team. New partnerships have expanded the network of organizations collaborating on AI safety, with a focus on areas like biosecurity and cybersecurity. The taskforce has also supported the establishment of the UK's AI Research Resource (AIRR), Isambard-AI, to conduct more intensive safety research. The report emphasizes the need to keep pace with rapidly advancing AI capabilities to address potential risks in fields such as cybersecurity and biosecurity. The taskforce aims to rigorously assess the harmful capabilities of new AI models and ensure safe development and deployment. It highlights the importance of state capacity in AI safety, with access to computing resources and collaborative efforts with leading AI companies. The upcoming AI Safety Summit will feature demonstrations on key risk areas, raising awareness and advocating for coordinated action in AI safety. The establishment of the AI Safety Institute signifies a long-term commitment to advancing AI safety for the public interest. [1]

[1] "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 Nov. 2023.

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 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. [1]

[1] "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence." The White House, The United States Government, 30 Oct. 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/.

1-2 November 2023. The AI Safety Summit 2023 is set to be hosted by the UK Prime Minister on November 1st and 2nd at Bletchley Park in Buckinghamshire, a location of great historical significance in the world of computer science and cryptography. This summit will serve as a crucial gathering point for international governments, leading AI companies, civil society groups, and research experts to deliberate on the challenges and risks associated with frontier AI development. While frontier AI models offer immense potential for economic growth and scientific advancement, they also bring about safety concerns that need to be addressed. The summit's objectives include establishing a shared understanding of these risks, fostering international collaboration for frontier AI safety, determining measures for organizations to enhance safety, identifying areas for research collaboration, and demonstrating how safe AI development can lead to its global, beneficial use. It reflects a concerted effort to promote responsible AI development and ensure its positive impact on society. [1]

[1] "Ai Safety Summit 2023." AI Safety Summit 2023 - GOV.UK, www.gov.uk/government/topical-events/ai-safety-summit-2023. Accessed 18 Oct. 2023.

5 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 like sampling and analysis for CBR agents, decontamination, waste management, biothreat contagion preparedness research, and more. [1]

[1] 2023 EPA International Decontamination Research and Development ..., *www.epa.gov/emergency-response-research/2023-epa-international-decontamination-research-and-development. Accessed 2 Nov. 2023.*

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. [1]

[1] Press, The Associated. "Europe Reaches a Deal on the World's First Comprehensive AI Rules." NPR, NPR, 9 Dec. 2023, www.npr.org/2023/12/09/1218374512/europe-first-comprehensive-ai-rules.

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 programme will bring the transformative power of rapid genetic sequencing technology and expertise of scientists to the forefront of our NHS." [1]

[1]
https://www.gov.uk/government/news/uk-and-us-announce-new-strategic-partnership-to-tackle-increased-biological-threats

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. [1]

[1] "A Pro-Innovation Approach to AI Regulation: Government Response." GOV.UK, www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response. Accessed 28 Mar. 2024.

18 March 2024. A gathering of prominent Western and Chinese artificial intelligence (AI) scientists took place in Beijing 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 guidelines and international cooperation in addressing the potential risks associated with advanced AI technology. [1]

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

22 March 2024. In a landmark move, the 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. [1]

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