ACMUN 2026

Disarmament and International Security Committee (DISEC)

BACKGROUND GUIDE

Agenda:

Global Surveillance of AI-Engineered Biothreats

EXECUTIVE BOARD

Nikhita Nandakumar - Co-Chairperson
Ryan Gomez - Co-Chairperson
Ruth Shetty - Moderator

Contents

Letter From the Executive Board

Dear Delegates,

Welcome to the Disarmament and International Security Committee at ACMUN 2026. We, the Executive Board, are delighted to have you join us in tackling one of the most urgent and complex issues of our time: the Global Surveillance of AI-Engineered Biothreats. As the convergence of artificial intelligence and synthetic biology accelerates, it brings both remarkable opportunities and profound risks to global security. This conference offers a unique platform to address these challenges head-on, and we are eager to see the innovative solutions you will bring to the table.

We are your Executive Board: Nikhita Nandakumar, serving as Co-Chairperson, and Ryan Gomez, serving as Co-Chairperson along with Ruth Shetty serving as your Moderator. Together we bring a shared passion for international security, technological ethics, and diplomatic collaboration. Our goal is to guide you through a rigorous and impactful debate that bridges science, policy, and diplomacy.

The rapid advancement of AI-driven tools like AlphaFold and RFdiffusion has transformed protein science, enabling unprecedented precision in designing novel biological structures. The significance of this shift was underscored when the 2024 Nobel Prize in Chemistry was awarded to the creators of these very tools. While these innovations promise breakthroughs in medicine and biotechnology, they also raise the spectre of misuse—whether intentional or accidental—in creating AI-engineered biothreats. Current international frameworks, such as the Biological Weapons Convention and UN Security Council Resolution 1540, provide a foundation for biosecurity, yet they were not designed to address the unique challenges posed by this technological convergence. National policies in leading countries like the United States, the European Union, and China are evolving, but they often treat AI and biotechnology as separate domains, leaving critical gaps in global surveillance and response.

Your task as delegates is to confront these gaps. We challenge you to develop robust, actionable frameworks for global surveillance systems that can detect and mitigate AI-engineered biothreats while balancing scientific progress with ethical and security considerations. Key questions to explore include: How can AI be integrated into early-warning biosurveillance networks without compromising privacy or sovereignty? What role should the United Nations and its bodies, such as the World Health Organization, play in coordinating these efforts? How can we prevent the misuse of emerging technologies by state or non-state actors?

To succeed, we urge you to dive into this background guide, conduct thorough research, and come prepared to propose concrete solutions. Your diverse perspectives and diplomatic skills will be essential in navigating this complex landscape. Let us approach this “whack-a-mole” challenge with adaptability and foresight, ensuring that our surveillance systems are not only reactive but also proactive and inclusive.

We are excited to witness your creativity, critical thinking, and dedication to shaping a safer world. Let ACMUN 2026 be a stepping stone toward meaningful global cooperation in the face of AI-engineered biothreats.

Warm regards,

Nikhita Nandakumar, Co-Chairperson

Ryan Gomez, Co-Chairperson

Ruth Shetty, Moderator

International Norms and UN Initiatives

Global governance of biotechnology and AI is framed by existing arms-control treaties and multilateral norms. The Biological Weapons Convention (BWC) explicitly prohibits the "development, production, and stockpiling of microbial or other biological agents" for hostile use, but it has historically lacked any standing verification or enforcement mechanism. Accordingly, the BWC has relied on confidence-building measures and occasional UN Security Council diplomacy rather than proactive surveillance. Complementing the BWC, UNSC Resolution 1540 (2004) requires all states to enact domestic laws preventing the proliferation of nuclear, chemical, or biological weapons—including to non-state actors. In practice, Resolution 1540 obliges states to control dual-use biotechnology exports and to cooperate on non-proliferation, but it does not itself establish new surveillance architectures.

Separately, the UN Secretary-General's Disarmament Agenda ("Securing Our Common Future") calls for strengthening global readiness to investigate alleged use of biological weapons. In particular, Action 10 of the 2018 Agenda directs states to improve "forensic" and epidemiological investigation capacity under the UN Secretary-General's Mechanism, although this is retrospective investigation rather than routine early warning.

A New Development: Verification Returns to the BWC

Following a mandate from the Ninth BWC Review Conference (2022), a Working Group on the Strengthening of the Convention has been meeting to address compliance and verification—the first sustained multilateral effort on BWC verification in roughly two decades. The Working Group held its sixth session in August 2025 and its seventh in December 2025, and convened an extraordinary eighth session in February 2026 to continue work on its final report. Proposals under discussion include establishing an Open-Ended Working Group to develop compliance and verification measures—potentially legally binding—covering on-site inspections, mandatory declarations, and investigative capabilities, drawing on lessons from the OPCW and IAEA regimes. For an agenda centred on surveillance, this process is directly relevant: it is the clearest sign yet that states are willing to revisit the BWC's long-standing verification gap.

The World Health Organization and the 2025 Pandemic Agreement

The World Health Organization also establishes international surveillance norms that affect biotechnology. The legally binding International Health Regulations (2005) require all WHO member states to develop core surveillance and reporting capacities for "any illness or medical condition" that may constitute a public health emergency. These IHR obligations create a global disease-reporting network—including WHO's Early Warning, Alert and Response System and regional bodies such as the European CDC—but they were devised for natural epidemics and do not explicitly anticipate engineered biothreats or AI-generated pathogens.

On 20 May 2025, the 78th World Health Assembly formally adopted the world's first Pandemic Agreement, following more than three years of negotiation prompted by the COVID-19 response. The agreement was approved in committee by a vote of 124 in favour, none against, and 11 abstentions before its adoption by consensus in plenary. It commits states to data-sharing, joint surveillance, and more equitable access to medical countermeasures. Importantly, the agreement is not yet in force: it requires a Pathogen Access and Benefit-Sharing (PABS) annex, after which it opens for signature and must be ratified by 60 states to enter into force. The 2024 targeted amendments to the International Health Regulations complement it. Even so, this framework remains focused on zoonotic outbreaks and equity rather than AI-enabled synthetic threats.

AI Governance at the UN

UN and UN-affiliated bodies have also begun to address AI governance in broad terms. In 2021, UNESCO secured unanimous adoption of a Recommendation on the Ethics of Artificial Intelligence, establishing

principles—human oversight, transparency, fairness—for AI worldwide. While not legally binding, it represents the first global AI standard and underscores the importance of human oversight of AI systems in alignment with human rights.

More recently, and building on the 2024 Global Digital Compact, the UN General Assembly adopted Resolution 79/325 on 26 August 2025, establishing the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on AI Governance. The 40-member Panel—appointed in 2026 and the first standing UN scientific body of its kind—will issue an annual, policy-relevant assessment of AI's opportunities and risks, while the Global Dialogue provides an intergovernmental forum for coordination. These mechanisms are general-purpose rather than bio-specific, but they offer delegates a ready-made institutional channel through which AI-biosecurity concerns could be raised. In sum, while the UN framework (BWC, IHR, Resolution 1540) provides legal norms against the misuse of biotechnology, and AI-ethics instruments set high-level guardrails, there is still no UN treaty that explicitly unifies biotechnology and AI in the context of bio-surveillance or early warning.

National Policy Frameworks

United States

U.S. policy treats biotechnology and AI through separate, extensive frameworks, with emerging links but few integrated mandates. On the biotechnology side, the National Biodefense Strategy (2022) and successive National Security Memoranda emphasise modernising bio-surveillance and improving early warning against biological threats. Most notably, on 5 May 2025 President Donald Trump issued an Executive Order on "Improving the Safety and Security of Biological Research" (EO 14292), which ends federal funding of dangerous gain-of-function research in "countries of concern," pauses certain high-risk domestic research pending a stronger oversight policy, and directs the Office of Science and Technology Policy (OSTP) and the National Security Advisor to develop enforcement mechanisms within 120 days. Unlike previous guidance, the order contains explicit enforcement and reporting provisions, representing the broadest tightening of U.S. biosecurity policy in decades.

In parallel, in May 2024 the White House OSTP consolidated and expanded federal guidelines on "Dual Use Research of Concern" (DURC) and "Potential Pandemic Pathogen" oversight, combining previous NIH frameworks and extending them with detailed review criteria and greater institutional responsibility. These directives cover a wide range of biotech research but focus on biological materials and experiments; they acknowledge "risks from computational approaches" for the first time but do not squarely regulate AI tools.

On AI governance, U.S. policy shifted sharply in 2025. President Biden's 2023 AI Executive Order (EO 14110), which had called for robust testing and risk management of AI systems and notably flagged biotechnology among the most pressing AI security risks, was rescinded by President Trump on 20 January 2025. It was replaced on 23 January 2025 by "Removing Barriers to American Leadership in Artificial Intelligence," which reorients federal policy toward AI dominance and deregulation and directed the drafting of a national AI Action Plan. (For accuracy, the earlier 2020 AI Executive Order, EO 13960, was issued by the first Trump administration, not by President Biden.)

That AI Action Plan—"Winning the Race: America's AI Action Plan"—was released in July 2025 and, significantly for this committee, treats biosecurity as a priority. It requires all institutions receiving federal research funding to use nucleic-acid synthesis providers with robust sequence-screening and customer-verification procedures, effectively converting previously voluntary screening standards into a funding condition, and it tasks OSTP with convening industry to build a data-sharing mechanism so providers can jointly flag fraudulent or malicious customers. This is the most concrete U.S. step yet toward AI-aware biosecurity screening, though it operates through funding conditions rather than binding statute. Overall, the United States retains extensive biodefense and AI policies that still largely run in parallel silos—now with a more deregulatory posture on AI but a tighter one on biosecurity.

European Union

The European Union is developing an ambitious regulatory agenda for both AI and biotechnology, with growing attention to security. On the AI front, the EU's Artificial Intelligence Act was formally adopted in 2024 and entered into force on 1 August 2024, classifying AI applications by risk. Its provisions apply in phases: prohibitions on "unacceptable-risk" uses and AI-literacy obligations took effect in February 2025; obligations for general-purpose AI models and the new EU AI Office became applicable in August 2025; and the bulk of the Act, including most high-risk-system rules, becomes generally applicable on 2 August 2026, with an extended transition to 2028 for AI embedded in regulated products. Notably, models used exclusively for scientific research are excluded from the Act's oversight, so AI tools used in biotech research (such as protein-design models) generally fall outside its remit unless deployed in healthcare or security contexts.

In parallel, the EU is restructuring its biotechnology framework. On 16 December 2025 the European Commission published its proposal for a European Biotech Act, a central pillar of the EU's life-sciences strategy intended to harmonise biotech rules across member states, streamline approvals, and align bio-innovation with security. Reporting indicates it is designed to integrate biosecurity considerations and to complement the existing Dual-Use Regulation (EU 2021/821), imposing new reporting obligations on research institutions for risky practices. The proposal is moving through the legislative process—with parliamentary and Council positions expected through 2026, trilogues possibly in the fourth quarter of 2026, and adoption not anticipated before the end of 2026—and a second package ("Biotech Act II") covering industrial, agri-food, and marine biotechnology is expected in autumn 2026.

Beyond these flagship laws, the General Data Protection Regulation (GDPR) and the European Health Data Space create a rights-based regime for health data that could limit open AI-driven analysis of genomic data, while the European Centre for Disease Prevention and Control (ECDC) coordinates infectious-disease surveillance under the IHR framework. The EU also invests in biodefence and genomic surveillance through HERA (the Health Emergency Preparedness and Response Authority) and Horizon Europe, though most funding is for diagnostics and countermeasures rather than AI-enabled threat forecasting. As with the United States, the EU's AI and biotechnology frameworks still operate in largely separate silos.

People's Republic of China

China has rapidly expanded its controls on both biotechnology and AI within its national-security framework. In 2020 it promulgated a comprehensive Biosecurity Law (effective 2021) that declares biosecurity a national-security priority, mandates national coordination, and imposes stringent controls on gene editing, human-subjects research, and pathogen handling. Crucially, Article 14 requires the government to "establish a biosecurity risk monitoring and early-warning system," directing the National Coordination Mechanism to develop national surveillance of biological risks—making China one of the few countries to legislate an overall biological-security intelligence system. The law also prohibits unapproved biotech activities that endanger national security and provides for criminal penalties for illicit genetic engineering.

On AI governance, China is likewise active. In 2023 it launched the Global AI Governance Initiative, outlining a vision of safe, "people-centric" AI development, and it was among the first major states to impose binding domestic AI rules—measures on algorithmic recommendation, deep synthesis, and generative AI that mandate security assessments and labelling of AI-generated content. These rules aim at social stability and content management rather than biotechnology. Analysts note that Beijing's AI policies stress sovereignty and leadership but do not directly address bioengineering risks.

In practice, China's approach is top-down and security-driven. Its National Security Law (2015) implicitly covers biotech under broad security mandates, and the Biosecurity Law's monitoring requirement suggests state-run biosurveillance via the China CDC and multi-agency task forces, alongside strict export controls on pathogens and gene sequences. Independent evaluations of how China operationalises AI in surveillance remain scarce, and there is no known directive requiring AI to scan specifically for engineered outbreaks. Overall, China's laws create unified national systems for biosecurity surveillance, but AI-bio convergence is left largely to strategic interpretation rather than detailed policy.

Convergence of AI and Synthetic Biology: Gaps and Emerging Focus

Across these international and national frameworks, a notable gap remains in addressing the convergence of AI and synthetic biology as a biothreat. Most governance instruments treat biotechnology and AI as separate domains. U.S. reforms on DURC and pandemic pathogens expand scrutiny of wet-lab experiments but say little about algorithmic tools that design pathogens. The EU's AI Act explicitly exempts scientific-research models, meaning cutting-edge "biological AI" used in labs escapes scrutiny under that law. Chinese AI rules focus on consumer and information security rather than synthetic biology. As a result, current policies generally fail to anticipate fully autonomous, AI-driven pathogen engineering, and oversight bodies have been criticised for lagging behind rapid innovation.

Experts argue that surveillance and threat detection will need new approaches. Reviews of the field call for adaptive, interactive early-warning systems spanning laboratory accidents to deliberate misuse, acknowledging that even legitimate use by scientists could lead to unexpected developments in an AI-enabled era. This "whack-a-mole" challenge implies that traditional surveillance (pathogen sequencing, symptom monitoring) must be augmented by AI tools that flag suspicious genetic sequences or unusual patterns. Some scholars recommend capability-based oversight rather than static pathogen lists, reflecting the fact that AI could create threats not found on any banned list. The U.S. gene-synthesis screening framework has begun to require DNA providers to use public databases plus intelligence inputs to flag dangerous sequences—a step toward AI-supported screening—and the 2025 AI Action Plan has since made such screening a condition of federal funding.

The most striking recent evidence that this gap is real came in October 2025, when researchers demonstrated—in a peer-reviewed study—that open-source AI protein-design tools could generate synthetic variants of a toxin capable of evading the very screening software that DNA-synthesis companies rely on (discussed as a case study below). The episode moved the AI-bio convergence problem from theoretical concern to documented vulnerability. Closing the remaining gap will likely require bridging the biosecurity and AI-policy communities: incorporating AI risk assessment into DURC reviews, updating disease-surveillance criteria to include anomalies suggestive of engineering, and exploring federated AI platforms that scan global genetic databases for engineered patterns. As of 2026, these ideas are being actively piloted and debated but have not yet coalesced into binding international policy.

Latest Innovations in Protein Science

The integration of artificial intelligence into protein science has transformed protein structure prediction and de novo protein design, accelerating biological research and expanding applications in medicine, biotechnology, and synthetic biology. This section examines three pioneering AI-driven technologies—AlphaFold, RFdiffusion, and Protein Message-Passing Neural Networks (MPNNs)—detailing their methodologies, applications, and the ethical considerations surrounding their dual-use potential. The significance of this field was formally recognised when the 2024 Nobel Prize in Chemistry was awarded to David Baker for computational protein design and to Demis Hassabis and John Jumper for AlphaFold's protein-structure prediction—the two lines of work that underpin much of this guide.

AlphaFold: A Paradigm Shift in Protein Structure Prediction

Development and Evolution

AlphaFold, developed by DeepMind, represents a landmark achievement in computational biology, addressing the decades-long challenge of predicting protein three-dimensional structures from amino-acid sequences. Introduced in 2018, it gained prominence at the CASP14 assessment in 2020, where AlphaFold 2 achieved a median Global Distance Test score rivalling experimental methods. The latest iteration, AlphaFold 3, was announced on 8 May 2024 by DeepMind and Isomorphic Labs, incorporating a novel "Pairformer" architecture and a diffusion module that refines atomic coordinates from an initial cloud of atoms. In November 2024 the AlphaFold 3 model code and weights were released for academic use, broadening access considerably.

Capabilities and Applications

AlphaFold 3 extends prediction beyond single proteins to protein–protein, protein–DNA, protein–RNA, and protein–ligand complexes, as well as post-translational modifications, achieving high accuracy across nearly all molecular types in the Protein Data Bank. These capabilities have profound implications for drug discovery, vaccine design, and enzyme engineering: by predicting binding sites for inhibitors, AlphaFold 3 facilitates targeted therapeutics, while its ability to stabilise epitopes enhances antigen presentation for vaccines. The AlphaFold Protein Structure Database, containing over 214 million predicted structures, provides open access that has been used by more than two million researchers across 190 countries.

Dual-Use Concerns

Despite its transformative potential, AlphaFold's precision raises significant dual-use concerns. The ability to predict protein structures accurately could, in principle, be exploited to design harmful proteins such as immune-evasive pathogen variants or stabilised toxins. To mitigate these risks, DeepMind has restricted commercial use of AlphaFold 3 and implemented access controls on its public server to limit high-risk predictions. Ongoing discussions, such as those at biosecurity summits hosted by the Institute for Protein Design, emphasise the need for voluntary guidelines and international surveillance to ensure responsible use.

RFdiffusion: Pioneering De Novo Protein Design

Methodology and Architecture

RFdiffusion, developed by the Baker Lab, is a generative AI model for de novo protein design, built on the RoseTTAFold structure-prediction network and fine-tuned for protein-backbone denoising using diffusion-based generative modelling. Unlike traditional methods that rely on predefined scaffolds, RFdiffusion initialises with random residue frames and iteratively denoises them over roughly 200 steps, guided by a mean-squared-error loss to ensure global frame continuity. Inspired by image-generation diffusion models

like DALL-E, it enables the creation of diverse protein topologies, including monomers, cyclic oligomers, and symmetric assemblies.

Applications and Experimental Validation

RFdiffusion has addressed a wide range of design challenges, including unconditional monomer generation, fold-conditioned design (e.g., TIM barrels with 42.5% in-silico success), symmetric oligomer design, and functional motif scaffolding. It has successfully designed binders for influenza hemagglutinin with a 19% experimental success rate and nanomolar affinities, validated by cryo-EM at 2.9 Å resolution, and has generated metal-binding proteins and enzyme active-site scaffolds confirmed by X-ray crystallography. Its open-source availability through platforms like ColabFold and GitHub has democratised protein engineering, and the release of RFdiffusion All-Atom—extending design to DNA, RNA, and small molecules—has further broadened its utility.

Biosecurity Considerations

The open-source nature of RFdiffusion, while fostering innovation, amplifies biosecurity risks: the same capabilities that enable rapid vaccine-scaffold development could be misused to engineer proteins with pathogenic or toxic properties. Discussions at successive AI-safety summits have highlighted the potential for AI-driven protein design to contribute to biological-weapon development, underscoring the need for secure access controls and international policy frameworks.

Protein Message-Passing Neural Networks (MPNNs)

Architecture and Functionality

Protein Message-Passing Neural Networks are a class of graph neural networks that model proteins as graphs, with amino acids as nodes and physicochemical interactions (hydrogen bonding, hydrophobic packing) as edges. This architecture enables MPNNs to predict how sequence variations affect protein stability, folding, binding affinity, and catalytic activity, facilitating rational protein design. A prominent implementation, ProteinMPNN (Baker Lab), uses a three-layer encoder-decoder architecture to predict amino-acid sequences for given protein backbones. Trained on high-resolution PDB structures, it achieves a native-sequence recovery rate of 52.4%, significantly outperforming traditional methods like Rosetta (32.9%), and is far faster—1.2 seconds per 100 residues on a single CPU, compared with 258.8 seconds for Rosetta.

Applications and Ethical Implications

ProteinMPNN excels at designing monomers, cyclic oligomers, protein nanoparticles, and protein–protein interfaces, with applications in drug development and industrial biotechnology; it has rescued previously failed designs by generating sequences with higher thermostability and improved crystallisation. Its integration with tools like RFdiffusion enables comprehensive in-silico design pipelines, where RFdiffusion generates novel backbones and ProteinMPNN designs corresponding sequences. Yet its ability to optimise protein function raises the same biosecurity concerns as AlphaFold and RFdiffusion: the potential to engineer proteins with enhanced catalytic activity or binding affinity could be misused, necessitating ethical guidelines and oversight. The convergence of these technologies has redefined biological research, offering unparalleled precision while underscoring the critical need for robust governance, secure access controls, and international collaboration.

Mapping Technologies to Governance Gaps

The following table maps key AI and biotechnology tools against the international frameworks that nominally cover them and the AI-specific gaps that remain.

Technology	Relevant Framework	Current Coverage	AI-Specific Gap
AlphaFold 3	Biological Weapons Convention (BWC)	Prohibits development and possession of biological weapons, including modified pathogens.	No explicit mention of AI-driven predictive tools or governance for access to structure-prediction models.
RFdiffusion	BWC + UNODA guidance	Normative coverage of synthetic biology under general prohibition clauses.	No verification protocols for AI-generated novel proteins or real-time monitoring of de novo design.
General AI models	UN Independent Scientific Panel on AI (2025)	Provides evidence-based assessment of AI risks; acknowledges dual-use concerns.	No operational mechanism to apply AI governance to biotech platforms in research or commercial use.
Protein MPNNs	UNSCR 1540 (WMD non-proliferation)	Requires states to prevent WMD access by non-state actors.	No legal language on AI models that optimise sequences for increased pathogenicity or drug resistance.
Sequence design tools	WHO “One Health” + IHR (2005)	Integrates animal, human, and environmental data for outbreak detection.	No integration of AI-based sequence-optimisation flags into the One Health surveillance pipeline.

Synthetic Biology and Biotechnology Advances

Molecular Docking

Molecular docking is a computational method used to predict how a small molecule (ligand) will bind to a target protein or nucleic acid. It is central to drug discovery, vaccine development, and synthetic biology, enabling researchers to screen thousands of compounds for optimal binding affinity, selectivity, and stability without exhaustive lab experiments. In the context of AI-engineered biothreats, however, the same approach could be misused to design highly potent toxins, enzyme inhibitors, or immune-evasive molecular agents with precision. With the integration of AI—platforms such as AlphaFold for structural prediction and Protein MPNNs for sequence optimisation—molecular-docking pipelines can be automated, drastically reducing the skill barrier. This creates a scenario in which a non-state actor with basic computational access could design and refine biologically active compounds with limited oversight.

CRISPR-Cas9

Scientific Breakthrough and Applications

CRISPR-Cas9, often called the "genetic scissors," is one of the most transformative tools in modern molecular biology. Originating from a bacterial immune-defence system, it allows scientists to precisely, efficiently, and relatively cheaply cut and edit DNA in plants, animals, microorganisms, and humans. Its ability to target specific genes with high precision has revolutionised agriculture (drought- and pest-resistant crops), medicine (potential cures for inherited disorders), and research (accelerating the understanding of gene function). CRISPR's promise extends to cancer therapy, antimicrobial-resistance solutions, and conservation biology.

Contested Origins and Patent Disputes

Although Emmanuelle Charpentier and Jennifer Doudna were awarded the 2020 Nobel Prize in Chemistry for CRISPR, its discovery has been mired in controversy. Competing claims from other researchers, such as Feng Zhang at the Broad Institute and Virginijus Šikšnys, have led to fierce intellectual-property battles. Allegations that key aspects of CRISPR's application to eukaryotic cells were built upon without due credit have fuelled accusations of idea appropriation. These disputes are more than academic; they shape who controls CRISPR patents, profits, and global access to the technology.

Why It Is a Security Concern

The contested origins of CRISPR point to a deeper governance problem: if a technology with such far-reaching consequences can emerge from opaque and disputed channels, public trust in scientific oversight erodes. Concentrating control of such a powerful technology in the hands of a few institutions or nations without transparent regulation risks inequitable access, misuse for unethical purposes (e.g., designer babies or bio-enhancement programmes), and potential dual-use in military or national-security contexts. International policy must therefore focus on surveillance of high-risk research through coordinated UN frameworks such as the BWC and the WHO R&D Blueprint; global licensing and registry systems for high-precision genome-editing tools; and ethical and legal oversight for AI-assisted molecular-design platforms to prevent uncontrolled proliferation of sensitive capabilities.

Case Studies

Case Study 1 — The Microsoft "Zero-Day" Biosecurity Study (2025)

In October 2025, a team led by Microsoft's chief scientific officer, Eric Horvitz, published a study in the journal Science titled "Strengthening nucleic-acid biosecurity screening against generative protein-design tools." Over roughly two years, and under strict confidentiality, the researchers used open-source AI protein-design tools to generate thousands of synthetic variants of known toxic proteins—redesigning them so they might retain their dangerous properties while evading the biosecurity screening software that DNA-synthesis companies use to flag dangerous orders. The experiment found that most of the redesigned sequences could slip past existing screening systems: a genuine "zero-day" vulnerability in global biosecurity infrastructure.

Crucially, the team practised responsible disclosure. Rather than publish the specific sequences, they worked confidentially for around ten months with screening providers, the International Biosecurity and Biosafety Initiative for Science (IBBIS), and government partners to develop "patches" that strengthened the screening tools before the findings were released. The episode is instructive on several levels: it demonstrates that the AI-bio threat is real and near-term rather than speculative; it shows that current screening regimes are technically defeatable by widely available tools; and it offers a model of how red-teaming, responsible disclosure, and public-private cooperation can close vulnerabilities without publishing a blueprint for misuse. It also raises hard questions for any surveillance framework: patches address known evasion techniques, but a static defence will always lag behind adaptive AI design—the "whack-a-mole" problem in concrete form.

Case Study 2 — KJ Muldoon and Personalised Gene-Editing Therapy (2024–25)

In 2024, doctors at the Children's Hospital of Philadelphia treated a newborn, KJ Muldoon, with a customised CRISPR base-editing therapy for CPS1 deficiency, a rare genetic disorder affecting roughly 1 in 1.3 million infants. The treatment, designed specifically for KJ's mutation, used base editing delivered in lipid nanoparticles to correct the defect in the liver. This marked the first-ever patient-specific gene-editing intervention, offering hope for thousands of rare genetic diseases too uncommon for large-scale pharmaceutical development.

While KJ's recovery demonstrates remarkable therapeutic potential, the case underscores critical biosafety concerns. The same rapid, precise, and customisable platform that corrected KJ's mutation could, in unregulated environments, be adapted to modify pathogens or the human germline. The ease of retargeting CRISPR instructions means such treatments could be repurposed without years of safety testing if accessed by untrained or malicious actors. The case highlights the urgent need for global surveillance, ethical oversight, and regulatory frameworks that let personalised medicine advance responsibly while minimising dual-use risks.

Case Study 3 — The CRISPR Twins and He Jiankui's Continued Work

In 2018, Chinese scientist He Jiankui claimed to have edited the genomes of twin embryos—later born as Lulu and Nana—to make them resistant to HIV by disabling the CCR5 receptor, a heritable alteration. The announcement sparked global outrage: the experiment was medically unjustified (safer methods to prevent HIV already exist), carried risks of off-target mutations and unknown long-term effects, and bypassed international safety and ethical standards. From a biosafety perspective, it illustrated the dangers of premature, unregulated human germline modification—and from a national-security standpoint, the risk that such capabilities could be exploited for population-level genetic manipulation, enhancement programmes, or destabilising "genetic-superiority" narratives.

He Jiankui served a prison sentence and was released in 2022. He has since re-established a laboratory and, as of 2026, publicly proposes new gene-editing research—now framed around Alzheimer's-protective genetic variants and rare diseases, and stating that no embryos would be brought to term. His return raises continuing concerns: public-health risks from off-target effects, ethical and social division over "designer" interventions, and the erosion of scientific oversight if a researcher can resume controversial work after minimal consequences. Internationally, it risks regulatory instability (a race to the bottom on safety standards), cross-border medical tourism to jurisdictions with looser enforcement, and geopolitical friction over dual-use enhancement capabilities.

Case Study 4 — Base-Editing of Human Embryos Reignites the "Designer Baby" Debate (2026)

In mid-2026, a research team at Columbia University reported editing the DNA of early-stage human embryos with a precision not previously achieved, using base editing to alter three disease-linked genes in order to test the safety and limits of rewriting DNA at the earliest stages of life. Unlike conventional CRISPR, which cuts the DNA strand, base editing rewrites individual genetic letters, reducing certain kinds of collateral damage. No embryos were brought to term, and researchers stressed that significant safety obstacles remain before viable human embryos could be edited safely.

The announcement reignited the "designer baby" debate. Some researchers hailed it as a technical milestone that could one day prevent devastating inherited diseases; others warned it edges society closer to heritable enhancement that bioethicists compare to modern eugenics. The controversy deepened over commercial involvement: reporting noted that firms screening IVF embryos have begun marketing predictive models for complex traits such as intelligence, and that commercial sponsors are backing further embryo-editing research—raising fears that experimental advances could fuel demand among wealthy patients well ahead of any ethical or regulatory consensus. For DISEC, the case is a reminder that the governance gap is not only about weapons: the same precision tools that raise bioweapon concerns also drive contested civilian applications, and any oversight regime must grapple with both.

Case Study 5 — PCR and "Separating the Scientist from the Science"

Polymerase Chain Reaction (PCR), invented by Kary Mullis in 1983, transformed molecular biology by enabling rapid, precise amplification of DNA. Mullis's eccentric personality and controversial public views—dismissing the link between HIV and AIDS, and rejecting climate science—have long prompted debate about how a scientist's judgment and motivations shape the trajectory and application of their work. The PCR story illustrates that innovation does not occur in a moral vacuum: unchecked ambition or misaligned objectives can influence the safety, oversight, and intended uses of a technology, even one as fundamental and beneficial as PCR.

Today, PCR is indispensable in diagnostics, forensics, genetic engineering, and synthetic biology. When integrated with AI-driven bioinformatics, it can be optimised to design highly specific primers, predict secondary structures, and model gene edits with precision. This same integration can be misused: to help recreate extinct or highly pathogenic viruses, engineer antibiotic-resistance genes, or synthesise artificial plasmids for gain-of-function work. Combined with automated synthesis platforms, AI-enhanced PCR reduces both the time and the technical barriers required for complex genetic engineering—precisely the lowering of barriers that this agenda seeks to surveil and govern.

Questions a Resolution Must Answer (QARMA)

What Is the Problem?

Scope and novelty. AI-engineered biothreats are novel pathogens or bioweapons created or enhanced through advanced AI-driven biotechnology. Such tools can design organisms beyond those found in nature, optimising traits like virulence, transmissibility, or population targeting. Experts warn that a single engineered release could cause a global catastrophe comparable to or worse than COVID-19. Unlike traditional bioweapons, which have historically required expensive, state-run programmes, AI may drastically shrink the technical barrier, enabling even non-experts or small teams to attempt creating dangerous pathogens.

Distinction from natural threats. AI-augmented threats differ in precision and unpredictability. Whereas natural outbreaks arise without specific design, AI tools could produce agents engineered for high lethality or vaccine resistance, and could optimise them to target specific genetic groups or geographies. An AI-designed agent might evade the biological constraints of natural strains and be tailored for stealth or potency—fundamentally different from legacy bioweapons and demanding new forms of detection.

Acceleration by AI. AI models can automate expert knowledge, speeding up the biothreat R&D cycle. Foundation models and emerging "AI-scientist" agents can suggest genetic sequences, predict molecular interactions, and troubleshoot experiments, standing in for human expertise. Recent evaluations indicate frontier models already provide meaningful "uplift" on laboratory and design tasks, and 2025–26 red-teaming studies—including one in which an unguarded biodesign agent generated over a thousand novel toxic protein candidates—suggest capability is rising faster than safeguards.

Gaps in surveillance and detection. Current systems are ill-suited to these threats. Most synthesis-screening regimes rely on static "select-agent" lists or simple sequence matching, which the 2025 Microsoft study showed an AI-designed pathogen could evade. Public-health surveillance focuses on recognised diseases and symptom clusters, leaving novel engineered outbreaks potentially undetected until too late. There is still no global mechanism specifically to monitor AI-assisted bioresearch, and—despite the verification process now underway—the BWC has no operational verification for new technologies.

Why Is It a Problem?

Global health and security implications. An AI-enabled biothreat could have catastrophic effects. Analysts warn that AI guidance on pathogen design could dramatically increase the potential for catastrophic outcomes by making lethal-agent creation more accessible. A successful engineered pandemic would overwhelm health systems, collapse economies, and destabilise societies; biologically targeted attacks could become a new form of warfare or terrorism.

Risk from non-state actors. Lowered barriers empower terrorists, criminal groups, and rogue scientists. Historically, lone attackers attempted bioterror with limited success due to complexity, but AI could change this by filling knowledge gaps and providing step-by-step guidance. The asymmetry is grave: small groups or individuals could wield strategic-level biological weapons once possible only for advanced states.

Regulatory loopholes and democratisation. Rapid diffusion of biotechnology exacerbates the danger. CRISPR kits, cloud bio-labs, and open-access gene databases are widely available, yet export controls tend to cover only known toxins and pathogens, not novel AI-designed sequences. Both AI and biotech are inherently dual-use and increasingly democratised; regulating AI development alone will not suffice without stronger biosecurity foundations. Cross-border DNA orders, unregulated cloud labs, and uneven screening could be exploited to develop AI-enhanced bioweapons unnoticed.

What Has Been Done So Far?

Treaties and norms. The main legal framework is the BWC, which bans the development of bioweapons but lacks specific language on AI or synthetic biology. Crucially, its Working Group on Strengthening the Convention has—across sessions in 2025 and an extraordinary session in February 2026—placed verification back on the agenda for the first time in around two decades, with proposals for an Open-Ended Working Group on legally binding compliance and verification measures. Other rules (the Geneva Protocol, Australia Group export lists, WHO's IHR) address biological threats generally but predate AI-driven science.

National and regional initiatives. The U.S. "America's AI Action Plan" (July 2025) explicitly prioritises biosecurity, making robust nucleic-acid synthesis screening and customer verification a condition of federal research funding and tasking OSTP with a provider data-sharing mechanism to flag malicious customers; the May 2025 executive order (EO 14292) tightened oversight of gain-of-function research. The EU's AI Act is in phased force and its Biotech Act was proposed in December 2025. WHO and its member states maintain global disease-surveillance networks (GOARN, ProMED), but these focus on natural outbreaks.

Private sector and academia. Leading AI labs conduct internal red-team assessments of their models' biosecurity risks and work with government safety institutes; the 2025 Microsoft–IBBIS collaboration to patch screening tools is a landmark example of responsible disclosure. Universities and think tanks (NTI, CSIS, CNAS, Johns Hopkins) publish analyses and recommendations, and genomic-synthesis firms collaborate on voluntary screening standards. At the same time, commercial competition may prioritise speed over safety, and DIY-biology communities continue with limited external oversight.

What Are the Possible Solutions?

Technical surveillance frameworks. Develop AI-enabled global monitoring: a standardised AI-driven DNA-sequence screening regime that replaces static blacklists and is resilient to the evasion techniques demonstrated in 2025; expanded worldwide genomic-sequencing and data-sharing networks feeding AI anomaly detection; cryptographic "watermarks" or metadata embedded in synthetic DNA to make orders traceable; predictive biosurveillance that continuously scans publications, genetic databases, and health reports for early signs of engineered outbreaks; and vetted cloud-lab protocols to close automation loopholes.

International cooperation and trust-building. Strengthen multilateral collaboration: shared threat intelligence, joint biodefence exercises, aligned safety standards, and transparency measures such as voluntary declaration of high-risk research. Regional partnerships (EU, ASEAN) might pool resources for AI-biosurveillance, and a whole-of-society approach—governments, industry, academia, civil society—would support governance, risk monitoring, and technical guardrails. Forums like WHO expert committees, the BWC Working Group, and the new UN Global Dialogue on AI Governance can serve as venues for consensus-building.

Legal and policy measures. Update treaties and regulations to cover AI bio-risks: revise the BWC to mention gene editing, synthetic biology, and AI-enabled research explicitly, with agreed definitions of prohibited activities; establish a multilateral protocol with mandatory licensing or export controls on high-risk AI biodesign tools; enforce universal DNA-synthesis screening (including offshore orders) with customer identity verification; and consider a global biosafety-and-biosecurity agency analogous to nuclear-safeguards bodies. Any framework should remain flexible as AI and biotech evolve.

Role of the UN and international bodies. Empower global institutions to coordinate responses: WHO could integrate AI-threat awareness into pandemic preparedness and revise the IHR to cover synthetic outbreaks; UNODA and the General Assembly could open deliberations linking AI and the BWC; and the

new UN Independent Scientific Panel on AI could be asked to assess AI-bio convergence specifically. Existing forums (BWC meetings, WHO conferences, G7/G20 health summits) should place AI-bio security on their agendas.

What Are the Implications of These Solutions?

Sovereignty and over-surveillance. Enhanced global surveillance may raise political and legal concerns. States might resist mandatory data-sharing or international lab inspections as infringements on sovereignty, and intrusive monitoring could be perceived as espionage. Any framework must include strict privacy protections and operate under agreed protocols; WHO data policies, for example, emphasise anonymisation and safeguards to protect privacy and avoid stigmatisation.

Ethical and societal issues. Data-sharing is critical for early warning but must not violate individual privacy or enable discrimination. Scientific communities will insist on preserving legitimate research and avoiding censorship, since overly broad restrictions could hamper beneficial innovation. Equity is a further concern: advanced surveillance tools and countermeasures should be accessible to developing countries, or solutions risk deepening global inequalities. Frameworks must uphold data protection and human rights, with an explicit ethical framework to maintain public trust.

Risk of escalation or misuse. Overly aggressive controls might drive adversaries toward more clandestine methods or trigger a biosecurity arms race. If AI-driven surveillance becomes widespread, malicious actors may invest in evasion techniques—exactly the dynamic the 2025 screening study exposed—or shift to subtler threats. There is also a strategic balance: too much transparency about vulnerabilities could itself become a security liability. Measures must be carefully calibrated to reduce biothreats without encouraging escalation or curtailing legitimate activity. No single solution is without trade-offs.

References and Further Reading

Primary sources and reporting used in preparing this guide. Delegates are encouraged to consult primary documents directly.

[1] Biological Weapons Convention (BWC) — UNODA. https://disarmament.unoda.org/biological-weapons/

[2] BWC Working Group on the Strengthening of the Convention, Sixth Session (2025) and Eighth Session (2026) — UNODA. https://meetings.unoda.org/bwc-

[3] "BWC Working Group Addresses Verification and Biosecurity," Global Biodefense (Aug 2025). https://globalbiodefense.com/2025/08/11/biological-weapons-convention-sixth-session-addresses-verification-and-biosecurity/

[4] UN Security Council Resolution 1540 (2004) — UNODA. https://disarmament.unoda.org/wmd/sc1540/

[5] "World Health Assembly adopts historic Pandemic Agreement" (20 May 2025) — WHO. https://www.who.int/news/item/20-05-2025-world-health-assembly-adopts-historic-pandemic-agreement-to-make-the-world-more-equitable-and-safer-from-future-pandemics

[6] International Health Regulations (2005, amended 2024) — WHO. https://www.who.int/health-topics/international-health-regulations

[7] UNESCO Recommendation on the Ethics of Artificial Intelligence (2021). https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence

[8] UN General Assembly Resolution A/RES/79/325 (26 Aug 2025): Independent International Scientific Panel on AI and Global Dialogue on AI Governance. https://docs.un.org/en/A/RES/79/325

[9] "Improving the Safety and Security of Biological Research," Executive Order 14292 (5 May 2025) — The White House / Federal Register. https://www.federalregister.gov/documents/2025/05/08/2025-08266/improving-the-safety-and-security-of-biological-research

[10] "Removing Barriers to American Leadership in Artificial Intelligence" (23 Jan 2025) — The White House. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/

[11] Executive Order 14110 (issued 2023, rescinded 20 Jan 2025) — overview. https://en.wikipedia.org/wiki/Executive_Order_14110

[12] "Winning the Race: America's AI Action Plan" (July 2025) — The White House. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

[13] "Biosecurity Guide to the AI Action Plan" — Johns Hopkins Center for Health Security. https://centerforhealthsecurity.org/our-work/aixbio/biosecurity-guide-to-the-ai-action-plan

[14] EU Artificial Intelligence Act — implementation timeline, European Commission. https://ai-act-service-desk.ec.europa.eu/en/ai-act/timeline/timeline-implementation-eu-ai-act

[15] European Biotech Act — European Commission (proposal, 16 Dec 2025). https://health.ec.europa.eu/biotechnology-0/european-biotech-act_en

[16] China Biosecurity Law (2020, effective 2021) — FAO/FAOLEX. https://faolex.fao.org/docs/pdf/chn198696E.pdf

[17] "China's AI Governance Initiative and Its Geopolitical Ambitions" — CIGI. https://www.cigionline.org/articles/chinas-ai-governance-initiative-and-its-geopolitical-ambitions/

[18] Horvitz, E. et al. "Strengthening nucleic-acid biosecurity screening against generative protein-design tools," Science (2 Oct 2025). https://pubmed.ncbi.nlm.nih.gov/41037625/

[19] "Microsoft says AI can create 'zero day' threats in biology," MIT Technology Review (2 Oct 2025). https://www.technologyreview.com/2025/10/02/1124767/microsoft-says-ai-can-create-zero-day-threats-in-biology/

[20] "New study sets a precedent for responsible AI and biosecurity" — IBBIS. https://ibbis.bio/new-study-sets-precedent-responsible-ai-biosecurity/

[21] Jumper, J. et al. "Accurate structure prediction of biomolecular interactions with AlphaFold 3," Nature 630 (2024). https://doi.org/10.1038/s41586-024-07487-w

[22] 2024 Nobel Prize in Chemistry (Baker; Hassabis and Jumper) — NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2024/press-release/

[23] Watson, J. L. et al. "De novo design of protein structure and function with RFdiffusion," Nature 620 (2023). https://doi.org/10.1038/s41586-023-06415-8

[24] Dauparas, J. et al. “Robust deep learning–based protein sequence design using ProteinMPNN,” Science 378 (2022). https://doi.org/10.1126/science.add2187

[25] Kolata, G. “Baby Is Healed with World’s First Personalized Gene-Editing Treatment,” The New York Times (15 May 2025). https://www.nytimes.com/2025/05/15/health/gene-editing-personalized-rare-disorders.html

[26] “New gene-editing breakthroughs are reigniting the debate around ‘designer babies’,” CNN (8 July 2026). https://www.cnn.com/2026/07/08/science/human-embryo-gene-editing

[27] “‘Edited’ human embryos reveal secrets of our development — and fuel ethical debate,” Nature (2026). https://www.nature.com/articles/d41586-026-02027-0

[28] “Controversial Chinese scientist He Jiankui proposes new gene-editing research” — Center for Genetics and Society (2026). https://www.geneticsandsociety.org/article/controversial-chinese-scientist-he-jiankui-proposes-new-gene-editing-research

[29] “The convergence of AI and synthetic biology: the looming deluge,” npj Biomedical Innovations (2025). https://www.nature.com/articles/s44385-025-00021-1

[30] “The whack-a-mole governance challenge for AI-enabled synthetic biology,” PMC (2024). https://pmc.ncbi.nlm.nih.gov/articles/PMC10933118/