Exploring the Ethics of AI as a Potential Weapon of Mass Destruction

Carter Dibsie[1], Marcus Johnson[2], Ladonia Prioleau[2], Leila Walden[2]*
Corresponding author: Leila Walden (leilawalden05@gmail.com)

[1]Department of Systems and Information Engineering, University of Virginia, Charlottesville, VA, USA

[2]Department of Electrical and Computer Engineering, Hampton University, Hampton, VA, USA

Abstract–Artificial Intelligence (AI) is increasingly integrated into scientific and technological domains where misuse can have severe consequences for human safety and global security. While AI is often promoted for accelerating discovery and improving efficiency, comparatively less attention has been devoted to the ethical risks that arise in high-consequence applications. This paper examines whether Artificial Intelligence can be ethically characterized as a potential Weapon of Mass Destruction (WMD).

This work contributes to an ethical risk framework that integrates dual-use analysis (the evaluation of technologies), an examination of its responsibility, and the application of harm minimization principles. This framework is applied to a specific ethical risk: the capacity of AI systems to lower barriers of entry for the development of Weapons of Mass Destruction by enabling actors without a high level of experience to generate toxic chemical compounds by utilizing easily accessible computing resources.

Rather than addressing AI's broader military or societal impacts, this study focuses on AI-assisted chemical design and autonomous decision-making systems that may accelerate proliferation and reduce meaningful human oversight. The methodology consists of a qualitative ethical analysis grounded in established technology ethical frameworks, specifically applied to documented AI-enabled chemical modeling and automation capabilities. This approach evaluates how intent and accountability shift when AI systems participate in a high-risk weapons development process.

The analysis demonstrates that while AI on its own is not a weapon, its ability to accelerate chemical weapon design, decentralize access to destructive capabilities, and diminish meaningful human control poses a significant risk due to it becoming an enabler of WMD-related risks. These findings lead to a need for actionable governance and design recommendations that aim to strengthen oversight, preserve accountability, and mitigate dual-use risks in highly consequential scientific domains.

Keywords—Artificial Intelligence, dual-use risk, chemical modeling, risk analysis, engineering, AI ethics, Weapons of Mass Destruction

I. Introduction

Artificial Intelligence's (AI) potential risks in high-stakes applications such as national security and public safety remain underexplored. Furthermore, as AI systems become more capable and autonomous, their use in critical and sensitive domains raises ethical concerns from a risk perspective that extend beyond purely technical considerations, affecting developers, governing agencies, research institutions, and the broader global public [1].

This work examines the ethical concern that AI has the potential to function as an enabler for Weapons of Mass Destruction (WMD) by lowering the barriers of entry to technological development [2]. In particular, this analysis focuses on AI-assisted chemical design, where emerging computational models have the potential to generate and evaluate chemical compounds without a high level of human expertise [3]. This introduces the possibility that both individuals and groups that lack the needed training could access tools that cause serious societal and environmental risks.

Given these risks, there is a growing need for a more comprehensive ethical framework that addresses the dual-use nature of AI technologies [4]. This paper argues that as AI continues to expand into more consequential domains, it is essential to evaluate not just how it can be of great benefit but also how it poses

serious risks given its potential to reshape destructive capabilities. Overall, this has serious implications for global security, public safety, and governance as a whole [3].

II. **AI as a Potential WMD Enabler**

The integration of Artificial Intelligence (AI) into the military and scientific spheres represents a profound paradigm shift, leading many scholars to categorize the technology as a primary enabler for Weapons of Mass Destruction (WMD) [5][3]. As machine learning capabilities advance, the convergence of AI with chemical, biological, radiological, and nuclear (CBRN) threats creates a dual-use dilemma of unprecedented scale [6]. The primary concern extends beyond the creation of autonomous hardware; it lies in how AI systematically lowers the barriers to entry for developing and deploying traditional WMDs while introducing novel risks associated with algorithmic speed and the erosion of human oversight [2][7].

One of the most immediate and acute threats exists at the intersection of AI and biological warfare. The development of biological weapons typically requires a high degree of "tacit knowledge," specialized, hands-on expertise often restricted to state-level laboratories. AI is effectively digitizing this expertise, particularly in the stages of agent optimization and cultivation. Large Language Models (LLMs) and advanced biochemical modeling tools can now suggest genetic modifications to enhance the virulence or vaccine-resistance of pathogens like *B. anthracis* [4]. Furthermore, generative AI can design "de novo" proteins or viral structures that do not exist in nature, potentially bypassing current diagnostic sensors and medical countermeasures [6]. By utilizing AI-driven "cloud labs" to monitor environmental variables in real-time, non-state actors can optimize the delicate process of cultivation, reducing the high failure rate previously associated with amateur biological synthesis [8].

In the realm of chemical warfare, AI acts as a massive accelerant for the discovery and synthesis of lethal agents. Traditionally, identifying a toxic compound stable enough for weaponization was a years-long research endeavor. However, modern AI tools have turned this into a matter of hours. In a landmark 2022 experiment, an AI designed for drug discovery was inverted to search for toxicity; in less than six hours, it generated 40,000 potential chemical warfare agents, including known nerve agents and entirely new, highly lethal molecules [4][6]. Beyond discovery, AI tools such as retrosynthesis software provide step-by-step instructions for manufacturing these agents using "dual-use" precursors, chemicals that are not currently on international watchlists, thereby allowing actors to bypass the monitoring frameworks of the Chemical Weapons Convention (CWC) [2][7].

Beyond the direct synthesis of biological and chemical agents, AI significantly lowers the logistical and operational hurdles associated with weaponization and dissemination. For a substance to be classified as a WMD, it must be successfully delivered to a target population in a way that maintains its toxicity or virulence. AI-driven computational fluid dynamics (CFD) can simulate how microscopic particles behave in specific urban microclimates, accounting for wind tunnels created by skyscrapers or ventilation patterns in subway systems [7] [9]. By automating these complex physics simulations, AI enables actors to identify the optimal release points and weather conditions required to maximize casualties, a task that previously required specialized atmospheric scientists and high-end supercomputing clusters [8].

The impact of AI on nuclear and radiological security is equally destabilizing, primarily through its integration into Command, Control, and Communication (NC3) structures. While AI cannot easily simplify the physical enrichment of fissile material, it introduces systemic risks to nuclear deterrence. AI-enhanced Intelligence, Surveillance, and Reconnaissance (ISR) can process vast amounts of satellite data to track mobile missile launchers or "silent" submarines, potentially compromising the "Second Strike" capability that maintains global stability [3]. Furthermore, as AI reduces the decision-making window from minutes to seconds, human commanders face immense pressure to delegate launch authority to algorithms to avoid being outpaced by an adversary's "hyper-war" capabilities [3][1]. This creates a high-risk environment for "flash-escalations" or accidental nuclear exchanges triggered by algorithmic drift or sensor error.

Furthermore, the integration of AI into cyber-physical security creates a new frontier for "virtual" mass destruction by targeting the industrial control systems (ICS) that manage critical infrastructure. AI-enhanced malware can autonomously identify and exploit "zero-day" vulnerabilities in power grids, water treatment facilities, or chemical plants at a speed that outpaces human defensive responses [8][7]. By inducing a catastrophic failure in a nuclear cooling system or chemical storage pressure valve, an AI-driven cyberattack can achieve the same level of environmental contamination and mass lethality as a physical strike [5]. This blurring of the line between digital and physical warfare suggests that AI does not just facilitate the creation of traditional weapons but transforms existing infrastructure into a latent weapon of mass destruction.

Ultimately, the classification of AI as a WMD enabler stems from its role as a force multiplier for catastrophic risk [5][2]. Whether through the accelerated synthesis of biochemical toxins or the destabilization of nuclear deterrence through hyper-fast decision-making, AI alters the global security landscape faster than international policy can adapt [3] [6]. The lack of robust, enforceable frameworks to govern these dual-use technologies means that tools intended for medical and scientific progress may inadvertently provide the mechanisms for large-scale destruction [8]. Addressing this threat requires a proactive reassessment of non-proliferation strategies, shifting the focus from the control of physical materials to the governance of the intangible, algorithmic blueprints that define modern warfare [5][1].

III. Objectives, Evaluation Criteria, and AI Risk Mitigation Alternatives

The primary objective is to evaluate AI's role in enabling high-risk technologies, particularly in the context of weapons development, given that its capabilities significantly increase the severity and scale of potential harm [5]. This will establish a key criterion when it comes to weighing the ethical responsibility and risk management of scaling the amount of human oversight.

The evaluation framework is based on several key criteria. The first is risk amplification, which assesses whether AI increases the scale, speed, or accessibility of harmful capabilities [8]. Second, dual-use potential evaluates the ease with which AI systems can be repurposed for military or weapons-related applications [6]. Third, autonomy & human oversight measures the degree of independent decision-making and presence for meaningful human control, which is essential for ensuring a high level of accountability in high-stakes environments [1].

Beyond those three criteria, other factors capture the overall broader risk. Accessibility and proliferation risk consider how widely available the technology is and its potential for misuse by potential actors [3]. Security & safeguards examines the effectiveness of protections against unauthorized use, such as fail-safes and system controls [4]. Finally, ethical and legal alignment ensures compliance from AI systems when it comes to international law and establishes ethical principles governing advanced technologies [7].

To address those concerns, there are several alternative approaches that can be implemented. Strengthening governance frameworks will provide the needed framework for regulatory oversight. Enforcing access controls limits the availability to help identify which entities get authorized access to the technology. Integrating design safeguards such as auditing mechanisms and having a constant human in the loop will help ensure AI systems operate within both the ethical and safety boundaries [8] [4].

IV. Ethical Analysis & Theories

The integration of artificial intelligence (AI) into the design and development of weapons of mass destruction (WMD) presents significant ethical concerns, particularly in relation to accessibility, autonomy, and accountability. One of the most pressing issues is AI's ability to lower the barrier to entry for harmful technologies, enabling individuals with limited expertise to contribute to the development of chemical or biological threats [8] [2]. This democratization of capability increases proliferation risk and challenges traditional models of security and control. Furthermore, AI systems can accelerate the speed and scale of weapons development, reducing the time required to design and deploy harmful agents [6].

Another critical concern is the reduction of meaningful human oversight. As AI systems become

more autonomous, particularly in weapon design and decision-making processes, responsibility becomes increasingly diffused [9] [1]. This raises questions about accountability, especially when harm results from AI-assisted actions. Scholars argue that removing humans from critical decision loops undermines ethical responsibility and increases the risk of unintended consequences [1]. Additionally, the opaque nature of many AI systems limits transparency, making it difficult to audit decisions or assign blame [7].

From an ethical theory perspective, several frameworks help evaluate these concerns. Utilitarianism assesses whether AI-enabled WMD development produces greater overall harm than benefit; given the scale of potential destruction, such technologies are largely viewed as ethically unjustifiable. Deontological ethics emphasizes the moral duty to prevent harm and uphold human dignity, suggesting that the development of systems that facilitate mass harm violates fundamental ethical obligations [5]. Virtue ethics focuses on the intentions and character of those developing AI, questioning whether enabling such capabilities aligns with responsible scientific conduct.

Additionally, dual-use ethics play a central role, recognizing that AI technologies designed for beneficial purposes can be repurposed for harmful applications [3]. This creates a moral obligation for developers and institutions to anticipate misuse and implement safeguards. Finally, principles of responsible innovation stress the importance of governance, transparency, and proactive risk mitigation in guiding AI development [4].

Overall, while AI itself is not inherently a weapon, its capacity to amplify harm, reduce oversight, and expand access to destructive capabilities positions it as a critical ethical concern in the context of WMD development.

V. **Recommendations**

Given the significant ethical and security risks associated with the use of artificial intelligence (AI) in the development of weapons of mass destruction (WMD), a multi-layered approach to governance, oversight, and responsible innovation is essential. First, there must be the establishment of robust regulatory frameworks and international agreements that specifically address AI-enabled weapons development. Similar to existing nonproliferation treaties, global cooperation is necessary to limit the misuse of AI technologies and ensure compliance with international humanitarian standards [6] [3]. These frameworks should include clear definitions of prohibited uses, enforcement mechanisms, and accountability structures.

Second, maintaining meaningful human oversight in all high-risk AI applications is critical. Fully autonomous systems, particularly those involved in lethal decision-making or weapons design, should be restricted or prohibited to ensure that ethical responsibility remains with human actors [1] [9]. Human-in-the-loop models should be implemented to preserve accountability and reduce the likelihood of unintended or escalatory outcomes.

Third, developers and institutions must adopt secure design principles and controlled access measures to prevent misuse. This includes limiting access to high-risk AI models, implementing usage monitoring systems, and incorporating fail-safes to detect and mitigate harmful outputs [8] [7]. Additionally, organizations should conduct rigorous risk assessments throughout the AI lifecycle to identify vulnerabilities and prevent exploitation.

Fourth, there should be a strong emphasis on ethical training and responsibility within the scientific and engineering community. Researchers and developers must be equipped with the tools to understand dual-use risks and make ethically informed decisions about their work [5]. Embedding ethics into STEM education and professional practice can help foster a culture of accountability and proactive harm prevention.

Finally, continuous monitoring, transparency, and interdisciplinary collaboration are necessary to adapt to the evolving nature of AI technologies. Open dialogue between governments, academia, and industry can support the development of standards and best practices while ensuring that emerging risks are addressed in a timely manner [4] [2].

Collectively, these recommendations aim to balance innovation with responsibility, ensuring that AI

advancements do not contribute to the proliferation or enhancement of weapons of mass destruction.

References

[1] S. M. Pedron and J. de A. da Cruz, “The future of wars: Artificial intelligence and lethal autonomous weapon systems,” *International Journal of Security Studies*, vol. 2, no. 1, 2020.

[2] M. Matsaberidze, “Artificial intelligence and a weapon of mass destruction,” *American Journal of Chemical and Biochemical Engineering*, vol. 8, no. 1, pp. 1–14, 2024.

[3] T. Zedelashvili and A. Guchua, “Artificial intelligence and weapons of mass destruction,” *Ante Portas – Security Studies*, no. 21, pp. 84–94, 2024.

[4] C. Caruso, “The risks of artificial intelligence in weapons design,” 2024.

[5] D. A. Kukuruznyak, “Can artificial intelligence be a weapon of mass destruction?,” 2023.

[6] E. Javorsky and H. Chaudhry, “Convergence: artificial intelligence and the new and old weapons of mass destruction,” *Bulletin of the Atomic Scientists*, 2023.

[7] B. Dresp-Langley, “The weaponization of artificial intelligence,” *Frontiers in Artificial Intelligence*, vol. 6, 2023.

[8] M. Brundage *et al*., “The malicious use of artificial intelligence: Forecasting, prevention, and mitigation,” 2018.

[9] P. Feldman, A. Dant, and A. Massey, “Integrating artificial intelligence into weapon systems,” 2019.