



DEFENSE DOSSIER

ISSUE 41

DECEMBER
2024



U.S., CHINA LOCKED IN AI ARMS RACE WHERE THERE ARE NO WINNERS

Paul Triolo

ARTIFICIAL INTELLIGENCE AND ITS INFLUENCE IN CHINESE MILITARY THOUGHT AND OPERATIONS

Larry M. Wortzel

A NEW AGE OF DECEPTION IN WARFARE

Rand Waltzman

THE DUAL USE DILEMMA IN MILITARY AI ADVANCEMENTS

Aiden Parker

HOW AI IS TURBOCHARGING DISINFORMATION

Sophia R. Turing



AMERICAN FOREIGN POLICY COUNCIL

Explaining the World. Empowering Policymakers.



DEFENSE DOSSIER

DECEMBER 2024 | ISSUE 41

- 1. From the Editors** 2
Ilan Berman and Richard M. Harrison
- 2. U.S., China Locked in AI Arms Race Where There are No Winners** 3
Fragmented Innovation and the Perils of zero-sum competition.
Paul Triolo
- 3. Artificial Intelligence and its Influence in Chinese Military Thought and Operations** 8
The PLA is struggling to balance political control and technological progress.
Larry M. Wortzel
- 4. A New Age Of Deception In Warfare** 12
Smart Systems can still be fooled. That's a real problem.
Rand Waltzman
- 5. The Dual-Use Dilemma in Military AI Advancements** 16
The Ethical and Strategic Challenges of Dual-Use Technology.
Aiden Parker
- 6. How AI is Turbocharging Disinformation** 18
From Deepfakes to Deception: AI's Threat to Truth.
Sophia R. Turing



LETTER FROM THE EDITORS

Welcome to the December 2024 issue of AFPC's *Defense Dossier*. As artificial intelligence reshapes industries, societies, and military capabilities around the globe, the United States and its allies must navigate the profound challenges and opportunities posed by this transformative technology. Nowhere is this more critical than in America's emerging strategic competition with China, where AI is becoming both a tool of innovation and a battleground for dominance.

In this issue, we delve into the multifaceted implications of AI for global security. Our contributors explore the U.S.-China AI arms race, the influence of AI in Chinese military doctrine, the dual-use dilemma in defense applications, and the growing threats posed by AI-driven disinformation and deception in warfare. Together, these analyses highlight the need for thoughtful governance, robust safeguards, and strategic foresight to harness AI's potential while mitigating its risks.

As AI continues to redefine the global landscape, understanding its impact on defense and security is more urgent than ever. We hope this edition provides valuable insights and actionable ideas for policymakers and practitioners alike.

All the best,

Ilan Berman
Chief Editor

Richard M. Harrison
Managing Editor



U.S., China Locked in AI Arms Race Where There are No Winners

Paul Triolo

Today, the United States and China are engaged in what many describe as an AI arms race, with each nation striving for technological dominance in artificial intelligence. While advancements in AI promise transformative benefits, the escalating rivalry risks undermining global stability, fragmenting innovation, and exacerbating geopolitical tensions. But by framing AI development as a zero-sum game, both nations could miss critical opportunities to address shared challenges like climate change, global inequality, and healthcare.

HOW WE GOT HERE

The seeds of the current U.S.-China AI rivalry were sown over the past two decades, as both nations recognized the strategic significance of emerging technologies. U.S. concerns about China's technological ambitions escalated during the late 2000s and 2010s. Events like the Google Aurora cyber intrusions and China's "Made in China 2025" initiative highlighted Beijing's focus on achieving technological self-sufficiency and leadership.

By 2017, China had formally codified its ambition to become the global AI leader by the year 2030 in a National AI Development Strategy.¹ While that strategy was aspirational rather than prescriptive, U.S. policymakers and analysts interpreted it as a direct challenge. Around the same time, breakthroughs in machine learning and neural networks, driven by advances in GPUs and massive datasets, began to transform AI from a niche field into a strategic priority.

The 2021 report from the U.S. National Security Commission on AI identified "chokepoint" technologies, such as advanced semiconductors, as critical leverage points.² This eventually led the Biden administration to issue a sweeping series of export controls and

other restrictions on China's access to key hardware and software, accelerating Beijing's push for indigenous innovation. In particular, large and complex export control packages released in October 2022,³ October 2023,⁴ and December 2024⁵ all targeted China's ability to access and develop domestically advanced hardware for training generative AI models, which by 2024 had become a major focal point of AI development. What began as a strategic competition has evolved into a full-fledged AI arms race, with far-reaching consequences for global cooperation and technological progress.

Commerce Secretary Gina Raimondo has stressed on multiple occasions⁶ that the goal of the controls was to prevent China from accessing the most sophisticated, cutting-edge AI chips, in order to prevent Chinese firms from being able to develop so-called frontier AI models. U.S. officials stress that one major goal of the technology restrictions levied by the Biden administration is "throwing sand in the gears" of Chinese companies in their ability to continue to develop and improve generative AI and other advanced AI models. Officials cite China's military civilian fusion program as likely to facilitate the transfer of advanced AI capabilities from private sector companies, where they are currently being developed, to military end users.

CURRENT DYNAMICS

As we look ahead to the new Trump administration, and track developments within the fast-paced AI industry, the competition appears set to continue. U.S.-China AI competition will unfold across multiple dimensions: technological, economic, military, and geopolitical. Both nations are vying to lead in key areas of AI development, from autonomous systems to generative AI. This rivalry is defined by several key dynamics:

Paul Triolo is a Partner for China and Technology Policy Lead at the Albright Stonebridge Group. He advises clients in technology, including across the AI stack, financial services, and other sectors as they navigate complex political and regulatory matters in the U.S., China, the European Union, India, and around the world. Over the past decade, he has written extensively on China's rise as a technology power, particularly with respect to AI, and is one of a handful of western observers who maintains close ties to China's AI sector and developing AI safety ecosystem.



Technological "Chokepoints." As noted, the U.S. has employed export controls to limit China's access to advanced semiconductors and AI-related hardware. These measures aim to slow China's progress in developing frontier AI systems, but have also pushed Beijing to accelerate its efforts to achieve self-sufficiency.⁷ This tit-for-tat strategy risks entrenching a bifurcated AI ecosystem.

Military Implications. Both nations view AI as a potential game-changer in defense and national security. Autonomous weapon systems, drone swarms, and AI-driven surveillance are areas of particular focus. The narrative that AI superiority will confer decisive military advantages fuels further competition, even as the broader implications remain uncertain.

Economic Competition. AI has become a cornerstone of economic policy for both nations, driving innovation in sectors such as healthcare, manufacturing, and fi-

countries, such as healthcare, concerns about access to personal data are likely to continue to limit levels of cooperation.

Geopolitical Tensions. Issues like Taiwan's critical role in semiconductor manufacturing are exacerbating the rivalry. Global foundry leader Taiwan Semiconductor Manufacturing Company (TSMC), for example, produces most of the world's advanced semiconductors, making it a strategic focal point in the AI arms race. Any conflict involving Taiwan could thus have catastrophic consequences for the global tech ecosystem. To the extent that the goal of U.S. technology policy is to prevent China from gaining the capability to develop advanced AI while the U.S. and its allies race ahead, a largely unexamined issue is how Beijing will react to this dynamic—especially given that, for the foreseeable future, Taiwan and TSMC will remain the epicenter of AI hardware manufacturing.⁸

THE RISKS OF A ZERO-SUM APPROACH

Given the current dynamics, framing AI as a zero-sum competition creates significant risks that extend beyond the U.S. and China. Yet most of these risks, and the sizeable collateral damage that could result, remain largely outside of the Washington discussion surrounding U.S.-China technology competition in general and U.S.-China AI competition in particular. They include:

A fragmentation of Innovation. A bifurcated AI ecosystem would force nations to duplicate research efforts, wasting resources and slowing

progress. The lack of collaboration would impede the development of global standards and best practices for AI governance. For example, U.S.-China discussions around AI during the Biden era have been minimal. The two sides have held just one bilateral discussion, and the topics were primarily centered on the use of AI in military command and control. In November of 2024,

“
By framing AI development as a zero-sum game, the U.S. and China risk missing critical opportunities to address shared global challenges.”

nance. However, the duplication of efforts and restricted knowledge sharing between the U.S. and China hinder global efficiency and innovation. The Biden administration has also taken new steps to control cross border transfers of data, for example, making collaboration around AI datasets and model training more difficult. Even in areas where collaboration would benefit both




during the last meeting between Presidents Biden and XI, the two sides reached an agreement barring the use of AI for nuclear command and control.⁹ However, the exact nature of the agreement remains unclear, as do enforcement mechanisms and other confidence building measures. It likewise remains unclear how the new Trump administration will carry through with the provisions of the agreement.

At the multilateral level, while Biden administration officials have tentatively supported limited Chinese participation in the important Bletchley Park Process, initiated last November by former British Prime Minister Rishi Sunak and designed to work toward a framework for global AI governance, it remains unclear how the new Trump White House will approach participation in this and other multilateral processes.

Global instability. The AI arms race intensifies geopolitical tensions, particularly around Taiwan. Escalating mistrust could lead to conflicts over access to critical technologies or preemptive actions against perceived threats. The zero-sum approach is particularly applicable in the case of Taiwan and TSMC, given the dominant position that TSMC occupies in the global supply chain for advanced AI optimized semiconductors. Should the Trump administration continue efforts to cut mainland Chinese semiconductor design firms off from using TSMC to manufacture their designs, this will raise the risk that Beijing will have fewer incentives to pursue a peaceful resolution of the Taiwan issue.¹⁰

Missed opportunities. AI has the potential to address pressing global challenges, such as climate change, future pandemics, and economic inequality. However, achieving these breakthroughs requires international collaboration, a pooling of resources, and the sharing of insights—an approach incompatible with zero-sum thinking. Global shortages of computational power are likely to continue, and with no sign of any collaboration between the U.S. in China in the AI sector in sight, there are huge opportunity costs to pursuing a zero-sum framing. Although researchers in both countries continue to collaborate, there will be increasing pressure to



Export controls on advanced AI hardware have pushed China to accelerate its efforts toward technological self-sufficiency.

reduce or end such interaction around AI-focused research, particularly in areas that are increasingly deemed to have national security implications, such as frontier models and agentic AI.

AI safety risks. Fragmented governance frameworks make it harder to address the ethical and safety challenges posed by AI. Rogue actors or nations may exploit these gaps, increasing the risk of misuse or unintended consequences. The Chinese government, particularly the Cyberspace Administration of China (CAC) and related standards bodies, has been proactive in developing regulations and a longer-term framework for regulating AI. The Chinese AI safety ecosystem is also well developed, with multiple research institutes pushing forward with novel ways to test models for safety and reliability. China is arguably ahead of the United States in terms of thinking through the complex issue around AI regulation, while leading U.S. companies are also researching AI governance approaches and collaborating with the budding U.S. AI safety ecosystem. Both sides could benefit from a robust dialogue on AI governance, despite differences around issues such as datasets, censorship, and data privacy.



LOOKING FORWARD

Hence, despite the competitive dynamics, there is a strong case for the U.S. and China to prioritize collaboration over confrontation in AI development. The stakes are simply too high to ignore the potential benefits of cooperation. Yet with U.S. officials justifying a range of technology controls on slowing China's ability to develop advanced AI models, any effort to move away from a more confrontational approach around AI development will be challenging.

Part of the difficulty lies in focus. U.S. efforts are focused on military end uses of AI, when the vast majority of use cases for the application of generative AI applications cut across civilian sectors where there could be potential areas of cooperation. While it will be difficult to carve out areas of agreement that would allow for greater cooperation, given the stakes, the cost of not attempting to do so are high.

Shared global challenges. AI is uniquely suited to tackle complex global issues like climate change, healthcare disparities, and resource scarcity. Collaborative efforts could accelerate progress in these areas while fostering goodwill between nations. Joint research initiatives and shared resources could drive breakthroughs faster than isolated efforts. By pooling talent, data, and computational power, the U.S. and China can unlock AI's full potential for helping solve some of the toughest problems facing both countries and the planet.

Trust-building measures. Establishing bilateral dialogues, joint projects, and shared governance frameworks can begin to slowly reduce tensions and foster transparency. These measures are essential for mitigating risks and building a foundation for long-term cooperation. The U.S.-China dialogue focused on military use of AI is an example that could be extended to other areas, but would need to be coupled with some flexibility around touchy questions such as U.S. export controls and other efforts to slow the ability of Chinese companies to develop advanced AI models and applications.

RECOMMENDED STEPS

To move beyond the current trajectory of rivalry, the following steps are recommended:

Reduce National Security Dominance Over AI Policy. Both nations should shift their focus from militarizing AI to emphasizing its civilian applications. This includes prioritizing investments in healthcare, education, and sustainability over defense-related AI projects.

Establish Bilateral AI Dialogues. Structured discussions on AI safety, governance, and ethical norms can build trust and prevent misinterpretations of each other's intentions.

Incentivize Collaborative Research. Governments should create funding mechanisms to encourage cross-border AI collaborations. A "CERN for AI" could serve as a model for pooling resources and expertise to address global challenges.

Invest in AI Safety and Misuse Mitigation. Both nations should prioritize efforts to detect and prevent the misuse of AI technologies, such as misinformation campaigns or malicious applications.

Promote Multilateral AI Governance. Establishing a global coalition that includes all major AI stakeholders, including China, is essential for setting ethical standards and ensuring the responsible use of AI.

The U.S.-China AI arms race represents a critical crossroads for technology and geopolitics. While competition can spur innovation, the current zero-sum framing of AI development threatens to fragment global collaboration, stall progress, and heighten the risk of conflict. Shifting the narrative from rivalry to cooperation is not just desirable but necessary.

ENDNOTES

1. Graham Webster et al., "China's Plan to 'Lead' in AI: Purpose, Prospects, and Problems," Digichina, August 1, 2017, <https://digichina.stanford.edu/>



- [work/chinas-plan-to-lead-in-ai-purpose-prospects-and-problems/](#).
2. Final Report of the National Security Commission on Artificial Intelligence, n.d., <https://reports.nscai.gov/final-report/>.
 3. U.S. Department of Commerce, Bureau of Industry and Security, "Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)," October 7, 2022, <https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file>.
 4. U.S. Department of Commerce, Bureau of Industry and Security, "Public information on export controls imposed on advanced computing and semiconductor manufacturing items to the People's Republic of China (PRC) in 2022 and 2023," December 5, 2024, <https://www.bis.doc.gov/index.php/about-bis/newsroom/2082>.
 5. U.S. Department of Commerce, Bureau of Industry and Security, "Commerce Strengthens Export Controls to Restrict China's Capability to Produce Advanced Semiconductors for Military Applications," December 2, 2024, <https://www.bis.gov/press-release/commerce-strengthens-export-controls-restrict-chinas-capability-produce-advanced>.
 6. Evelyn Cheng, "U.S. export controls need to 'change constantly' even if it's tough for businesses, Secretary Raimondo says," CNBC, December 5, 2023, <https://www.cnbc.com/2023/12/05/commerce-sec-raimondo-us-export-controls-need-to-change-constantly.html>.
 7. Paul Triolo, "The Evolution of China's Semiconductor Industry under U.S. Export Controls," American Affairs, November 2024, <https://americanaffairsjournal.org/2024/11/the-evolution-of-chinas-semiconductor-industry-under-u-s-export-controls/>.
 8. For more on Taiwan issues, see Paul Triolo, "The Taiwan Debate Is Heading in a Dangerous Direction," The Wire China, July 14, 2024, <https://www.thewirechina.com/2024/07/14/the-taiwan-debate-is-heading-in-a-dangerous-direction-chip-industry/>.
 9. Jarrett Renshaw and Trevor Hunnicutt, "Biden, Xi agree that humans, not AI, should control nuclear arms," Reuters, November 16, 2024, <https://www.reuters.com/world/biden-xi-agreed-that-humans-not-ai-should-control-nuclear-weapons-white-house-2024-11-16/>.
 10. For more, see "Chapter 7: Sino-American Technology Competition and the Asia-Pacific," in Asia-Pacific Regional Security Assessment 2022 (London, IISS 2022), <https://www.iiss.org/publications/strategic-dossiers/asia-pacific-regional-security-assessment-2022/aprsa-chapter-7/>.



Artificial Intelligence and its Influence in Chinese Military Thought and Operations

Larry M. Wortzel

China's military leaders are concerned that it has been decades since the Chinese People's Liberation Army (PLA) has seen any combat—and that its tactics, equipment, and operational art date back to the 1960s. They are consequently in a search for new technologies they hope will improve the performance of the Chinese armed forces. These new technologies include artificial intelligence (AI) and rapid data management.

When Xi Jinping, the General Secretary of the Chinese Communist Party (CCP) and Chairman of the country's Central Military Commission (CMC), took office in 2012, he commented that PLA generals could not make decisions and did not understand the modern battlefield. He has complained about this throughout his tenure, and the senior ranks of the PLA give him no confidence that the military can perform its missions. Xi has had to fire CMC members, ministers of defense, and most recently the head of the PLA's political work department, all for corruption.

A LONGSTANDING QUEST

Even before Xi's critique, however, military thinkers in China were studying foreign wars and foreign militaries, hoping to get ideas that would help the PLA in its decision-making and stimulate new tactics. Progress in how the PLA commanders make decisions is hampered by a fear of not awaiting approval from the CCP's political commissars, who share decision responsibilities with unit commanders—a state of affairs that can lead to conflicts which slow decision-making or paralyze leadership.¹

The main model studied by the PLA for how a modern military should be equipped, manage command and control, and operate on the battlefield has been the U.S. military.² In turn, many PLA strategists and senior lead-

ers seem to believe, even hope, that artificial intelligence (AI) and large model data management can lead to better and faster decision making. Thus, the focus of much of the recent writing and research in the PLA is on ways that AI can effectively coordinate manned and unmanned combat systems in modern warfare.

For instance, a recent issue of the PLA's premier strategy journal, *China Military Science*, devoted about a third of its articles to exploring difference forms of AI controlled “kill chains” in combat operations and analyzed in detail how U.S. forces use AI.³ In the same vein, a recent article by a PLA officer from a unit that provides information support to military operations outlined the difficulties of transitioning from cyber warfare to artificial intelligence-supported “mosaic warfare.”⁴ That concept is one explored by the Pentagon's Defense Advanced Research and Development Agency (DARPA), and designed to conduct attacks in “in parallel across a wide front employing distributed sense, decide, and act systems.” It employs artificial intelligence directing a large number of platforms and weapons across several domains (air, land, sea, and space)⁵—an idea that the PLA has been seeking to emulate. Indeed, inside the PLA, recent thinking has evolved from “network-centric operations” to “decision-centric warfare” aided by artificial intelligence and unmanned systems with sensors to find and detect targets.⁶

However, implementing AI-enabled warfare faces significant hurdles within the PLA's institutional culture. Scholar Chen Yanbiao, writing in a volume for the PLA Academy of Military Science (AMS), argues that while Chinese military leaders understand the need for 21st century command and control methods, translating this understanding into practice remains challenging.⁷ In his book *Innovation in Intelligent Command Methods for Combat Operations in the Information Age*—which is part of the

Larry M. Wortzel, Ph.D., is a retired Army colonel who spent two tours of duty as a military attaché in China. He served as China Foreign Area Officer and a strategist. Dr. Wortzel also served for 19 years on the U.S.-China Economic and Security Review Commission. Presently he is Senior fellow in Asian Security at the American Foreign Policy Council.



modern PLA's "library on warfare"—Chen draws a telling historical parallel. Recalling the 1948 Liaoning-Shenyang campaign, when Communist forces achieved their first major victory against the Nationalists, he notes that senior leaders did not fully grasp operational dynamics while subordinates simply knew to "charge the gunfire." This gap between leadership understanding and battlefield reality persists today.

Chen characterizes PLA forces as "somewhat backward" when compared to the United States, particularly in developing the integrated, networked, and interactive systems required for information-age warfare.⁸ His analysis tracks the evolution of U.S. command and control systems from the 1970s to today, emphasizing how networked integration of reconnaissance, intelligence, and surveillance has produced a transformed "kill chain" on the battlefield (C4ISR). While the PLA dedicates substantial military and civilian resources to developing AI algorithms and data management, questions remain about effective implementation.⁹

GROWING PAINS

Recent PLA scholarship reflects this tension between ambition and reality. Officers from a Beijing-based unit focused on developing AI command systems published a comparative study in *China Military Science* examining command and control of unmanned forces in foreign militaries.¹⁰ Their article argues that blending manned and unmanned systems through AI has proven more effective than traditional approaches. The 2024 *Analysis of Typical Cases of UAV Combat Application* reinforces these findings, while AMS researcher Guo Zhi contends that

The PLA's dual command structure, where political commissars share authority with commanders, may hinder rapid AI-enabled decision-making.

future wars will require hybrid AI-control centered cognitive decision making.¹¹

Yet, even as some strategists push for technological innovation, others grapple with the ethical implications of AI-governed warfare. This debate reached its pinnacle during the November 2024 Asia-Pacific Economic Cooperation Summit in Peru, where Xi and President Biden discussed AI's role in nuclear weapons employment. According to a White House spokesman, "The two leaders affirmed the need to maintain human control over the decision to use nuclear weapons" and "stressed the need to consider carefully the potential risks and develop AI technology in the military field in a prudent and responsible manner."¹²

These promises, however, ring hollow on two counts. First, they focus primarily on strategic ICBM strikes, leaving open the question of AI's role in conventional warfare. Second, even current decisions on strategic strikes rely heavily on AI-sorted large data models for target identification.

And the influence of AI on the speed of war and the types of weapons to be used is a central issue for the PLA. Xi Jinping's has pointed out in his guidance to the PLA that in combat, intelligent technology (AI), unmanned equipment, and automated data sharing, are part of multi-domain, joint, distributed target chains. They are part of new combat trends.¹³

POLITICS AND TECHNOLOGY

The Chinese military's drive toward AI integration reveals several paradoxes in PLA modernization efforts. First, while Chinese military leaders recognize the need for faster, AI-enabled decision-making, the PLA's dual command structure—where political commissars share authority with unit commanders—may actually become more problematic as AI systems increasingly demand rapid, decentralized responses. Chen's observation that PLA forces remain "somewhat backward" compared to the United States reflects not just technological gaps, but deeper institutional challenges.

Second, the PLA appears to be pursuing contradictory approaches to AI integration. At the strategic nuclear level, as evidenced



by the Xi-Biden discussions in Peru, China signals restraint and human control. Yet at operational and tactical levels, PLA writings reveal an aggressive push toward AI-enabled "kill chains" and autonomous systems. This bifurcation may create dangerous uncertainties in crisis situations, particularly if adversaries misread which domains are governed by human decision-making versus AI systems.

The trajectory of PLA AI integration suggests several implications for U.S. defense planning. While Chinese military journals demonstrate a sophisticated understanding of AI's potential in modern warfare, translating this theoretical knowledge into practical capabilities requires overcoming significant organizational and cultural barriers. The PLA's continued study of U.S. military AI applications indicates both admiration for and concern about American advantages in this domain.

In the future, AI will indeed be a major part of military decision-making and warfighting for both China and the United States. However, the PLA's distinct institutional characteristics—its dual command structure, emphasis on political control, and relatively limited combat experience—suggest that its integration of AI may follow a uniquely Chinese path rather than simply mimicking U.S. approaches. Understanding these differences, and their implications for military competition in the Indo-Pacific, will be crucial for U.S. strategic planning in the decades ahead.

ENDNOTES

1. Larry M. Wortzel, *The PLA and Mission Command: Is the Party Control System Too Rigid for Its Adaptation by China?*, Association of the U.S. Army, Land Warfare Paper 159, March 14, 2024, <https://www.ausa.org/publications/pla-and-mission-command-party-control-system-too-rigid-its-adaptation-china>
2. The major sources for this article are: Chen Yanbiao, *信息时代作战智指挥方法创新* *Innovation in Intelligent Command Methods for Combat*

China's military journals reveal a sophisticated understanding of AI's potential but highlight significant organizational and cultural barriers to implementation.

Operations in the Information Age (Beijing: Military Science Press, 2020); Hu Yongjiang, Wang Zhili, eds., *无人机作战运用典型例分析 Analysis of Typical Cases of UAV Combat Application* (Beijing: Weapons Industry Press, 2024); Jing Tao, “基于联合作战体系‘有人 + 无人’协同作战深化研究 In-depth research on ‘manned + unmanned’ collaborative operations based on joint combat systems,” *Military Arts* 1 (2022), 3-6; Zhang Kaiyue and Zhang Huang, “军用人工只能治理：风险，困境，于中国智慧 Military Artificial Intelligence Governance: Risks, Dilemmas, and Chinese Solutions” *China Military Science* 4 (2023), 67-79; Zhang Huang, Lin Han, and Zhang Kaiyue, “智能化战争的伦理风险治理路径 Ethical Risks and Governance Paths of Intelligent Warfare,” *China Military Science* 5 (2022), 109-117; Huang Jianming and Liu Kui, *论指挥决策 On Command Decision-Making*, *China Military Science* 2 (2022), 125-131; Lin Xiangyang, et.al., *外军无人坐镇力量编组与指挥控制 Discussion on Formation and Command and Control of Unmanned Combat Forces in Foreign Armies*, *China Military Science* 5 (2023), 128-132; Guo Zhi, “未来战争将是认知决策为中心的混合型控制站 Future War Will be a Hybrid War of Control Centered on cognitive Decision Making: Research and Thinking on Future Intelligent Warfare,” *China*



- Military Science* 1 (2024), 1-8; Zhang Song and Yang Liqun, eds., *网上指挥对抗训练创新研究 Innovative Research on Online Command Confrontation Training* (Beijing, Hai Hu Publishing House, 2016); Yang Cunyin, “从“赛博战”到“马赛克战” From Cyber Warfare to Mosaic Warfare,” *PLA Daily*, September 14, 2021, 7, http://www.81.cn/jfjbmap/content/1/2021-09/14/07/2021091407_.pdf.pdf;
- Liu Haijiang, Zhang Shanshan, “美军马赛克的制胜机理关键技术及作战运用 Winning Mechanisms, Key Technology, and the Operational Application of the U.S. Military’s Mosaic Warfare,” *China Military Science* 1 (2022), 140-147.
3. *China Military Science* 2, iss. 194 (2024), 47-54, 55-63, 72-79, 135-144.
 4. Yang, “From Cyber Warfare to Mosaic Warfare.”
 5. See Department of Defense, DARPA, “DARPA Tiles Together a Vision of Mosaic Warfare,” n.d., <https://www.darpa.mil/work-with-us/darpa-tiles-together-a-vision-of-mosaic-warfare>; See also David A. Deptula and Heather Penney, “Mosaic Warfare,” *Air and Space Forces*, November 1, 2019, <https://www.airandspaceforces.com/article/mosaic-warfare/>.
 6. See Guo, “Future War Will be a Hybrid War of Control Centered on cognitive Decision Making: Research and Thinking on Future Intelligent Warfare.”
 7. Chen, *Innovation in Intelligent Command Methods for Combat Operations in the Information Age*, 21-23, 40-41.
 8. Ibid., 40-41.
 9. Jeffrey Ding, “China’s AI Implementation Gap,” *China Leadership Monitor* iss. 82 (2024), <https://www.prcleader.org/post/china-s-ai-implementation-gap>.
 10. Lin Xiangyang et.al., “Discussion on Formation and Command and Control of Unmanned Combat Forces in Foreign Armies,” 128-132.
 11. Hu and Wangi, eds., *Analysis of Typical Cases of UAV Combat Application*.
 12. Lauren Egan and Phelim Kine, “Biden’s final meeting with Xi Jinping reaps agreement on AI and nukes,” *Politico*, November 16, 2024, <https://www.politico.com/news/2024/11/16/biden-xi-jinping-ai-00190025>
 13. “习近平谈人工智能：赢得全球科技竞争主动权的重要战略抓手 Xi Jinping talks about artificial intelligence: an important strategic grasp to win the initiative in global scientific and technological competition” <http://cpc.people.com.cn/xuexi/n1/2018/1101/c385476-30376558.html>; 习近平在视察空降兵军时强调 全面加强练兵备战 提高空降作战能力 努力建设一支强大的现代化空降兵部队 When inspecting the Airborne Troops, Xi Jinping stressed the need to comprehensively strengthen military training and preparation, improve airborne combat capabilities, and strive to build a powerful and modern airborne troops,” *Observer*, November 5, 2024, https://www.guancha.cn/military-affairs/2024_11_05_754269.shtml



A New Age Of Deception In Warfare

Rand Waltzman

A ground-based missile defense system using machine learning to identify and intercept incoming threats is deployed to protect a strategic base. The system relies on radar signatures to classify objects as friendly, hostile, or non-threatening. Adversaries launch a coordinated attack by releasing a swarm of small drones programmed to emit radar signals that mimic incoming missiles. The defense system, overwhelmed by the sheer number of targets, prioritizes the drones, expending its missile interceptors. Amid the confusion, a genuine missile strike penetrates the base's defenses, causing significant damage.

While the scenario above is fictional, it represents a plausible future scenario. It helps to highlight how even sophisticated systems can be overwhelmed and rendered ineffective through deliberate deception, and underscores the critical need for robust validation mechanisms and human oversight. As we stand on the brink of a new technological era, the integration of machine learning and artificial intelligence into military systems introduces not only promise but also profound risks.

Behind this realization is a stark reality. Deception, an ancient tactic in warfare, is evolving. Where humans were once the primary targets of such operations, the rise of intelligent machines opens a new battlefield: deceiving technology itself. Unlike traditional cyberattacks, which often rely on breaching a system's defenses to access its internal data or software, deceiving machine learning systems requires no such intrusion. Instead, adversaries can manipulate these systems by exposing them to carefully crafted inputs through external sensors, such as cameras, microphones, or radar systems.

This approach mimics the way humans can be deceived—not by tampering with the brain directly but by presenting misleading or false information to the senses. For example, altering a stop sign's appearance to confuse

an autonomous vehicle into misidentifying it as a speed limit sign relies entirely on manipulating what the vehicle "sees." This kind of attack exploits the system's reliance on patterns and inputs, demonstrating how deception has transcended its historical focus on human perception to target the very tools designed to assist us.

LEARNING BY EXAMPLE

The ability to learn by example is one of the most powerful and mysterious forces driving intelligence, both in humans and machines. Think of how children learn to recognize a letter or an animal. By being exposed to enough examples—whether it's the letter "B" or a cat—they effortlessly develop the ability to identify new instances of that letter or animal. This seemingly magical process stems from the brain's natural tendency to find patterns. It identifies the essential qualities of "cat-ness" or "B-ness," allowing the child to categorize experiences unconsciously and automatically. However, while this process feels intuitive, explaining it remains one of the great mysteries of intelligence. The issue isn't forgetting which examples contributed to learning; it's losing track entirely of which inputs shaped the knowledge in the first place. What remains are judgments, disconnected from their origins, forming the bedrock of human cognition without a clear roadmap of how they were constructed.

Deep learning, one of the most transformative forms of artificial intelligence, operates much like human cognition because it was inspired by this very mechanism of learning by example. Just as children learn through exposure, deep learning systems are trained using labeled examples of what they are expected to recognize. Over time, these systems develop neural networks capable of identifying new, unseen data. For instance, a machine shown

Rand Waltzman has over 40 years of experience working in AI. This includes two terms as an AI Program Manager at DARPA. The first term managing the DARPA Image Understanding Program ended in 1991. The second term managing the Social Media in Strategic Communications program and the Anomaly Detection at Multiple Scales insider threat detection program ended in 2015. He is currently Adjunct Senior Information Scientist at the RAND Corporation studying the next generation of manipulation, influence and deception in Virtual and Augmented Reality environments.



enough pictures of cats can learn to recognize cats it has never encountered before. Yet, just like human intelligence, the process by which deep learning systems make decisions is shrouded in opacity. Once trained, the system retains no record of the inputs it learned from or how those inputs informed its decision-making. This lack of transparency in both human and machine cognition underscores a shared trait: an inability to explain the logic behind decisions.

This opacity in deep learning is known as the “black box problem,” and it presents significant challenges. The fundamental issue lies in our inability to understand how a system arrives at its conclusions. This becomes particularly problematic when deep learning systems produce undesirable outcomes. For example, if an autonomous vehicle fails to stop for a pedestrian and causes an accident, diagnosing the error is an immense challenge. Did the system misinterpret the pedestrian’s presence? Was the situation too novel for the system to handle? Without insight into the system’s “thought process,” identifying and addressing the root cause becomes nearly impossible.

Even if retraining the system with new scenarios could improve performance, the complexity of real-world conditions makes it impossible to anticipate every potential one. Consider scenarios such as sunny weather with light fog, salt-streaked roads after a winter storm, or countless other variations. These infinite permutations make it exceedingly difficult to ensure that a system has seen and learned from every possible condition it might encounter. As a result, questions about the robustness and reliability of deep learning systems remain unresolved. But it is worse than that.

THE DAWN OF MACHINE-TARGETED DECEPTION

Deception has always played a critical role in warfare, from feigned retreats in ancient battles to disinformation campaigns during the Cold War. But as warfare becomes increasingly reliant on AI and machine learning, malicious actors are adapting their strategies to exploit these systems.

Rapid advances in machine learning have brought about technologies capable of analyzing vast amounts of data, classifying information, and even making decisions

**Deception, an ancient tactic in warfare,
is evolving to target the intelligent
machines we rely on.**

autonomously. In military applications, these technologies could be embedded in decision support systems, targeting algorithms, and logistics planning. Yet, as reliance on machine learning grows, so do the associated vulnerabilities. Unlike humans, machine learning systems can be manipulated through carefully crafted inputs. These inputs may deceive the system into misinterpreting its environment, leading to catastrophic outcomes—like friendly fire or misdirected troop deployments.

Machine learning systems, while transformative, remain vulnerable to creative forms of manipulation that exploit their reliance on external inputs. For example, facial recognition systems used for security can be deceived by adversaries wearing carefully crafted masks or applying specific patterns of makeup. These subtle changes can confuse the system into misidentifying individuals or failing to recognize them entirely, compromising secure facilities or sensitive operations. Autonomous drones that rely on GPS signals for navigation can be misled by spoofed location data. An adversary could trick the drone into believing it is on course while redirecting it to a location where it could be captured or destroyed. Natural language processing (NLP) systems, which generate or analyze human-like text, are similarly susceptible to manipulation.

Operational decision-making algorithms, such as those used in route planning, are another vulnerable target. By injecting false traffic or weather data, adversaries could trick the system into recommending inefficient or even hazardous routes, delaying the transport of troops or supplies in a combat zone. Predictive maintenance systems, which use machine learning to identify equipment failures before they occur, could be manipulated by falsifying sensor data. This might lead to premature equipment



Unlike traditional cyberattacks, deceiving machine learning systems requires no intrusion—only carefully crafted inputs.

replacements, unnecessary downtime or, worse, failures at critical moments. Environmental monitoring systems, designed to detect chemical or biological threats, could be deceived through the introduction of benign substances that mimic the signatures of dangerous agents, prompting unnecessary evacuations or diverting resources away from actual threats. These diverse examples illustrate how adversaries can exploit machine learning systems, not by breaking into them directly but by subtly corrupting the information these systems rely on, making deception a widespread and versatile threat.

Compounding these risks is the opaque nature of machine learning. While engineers can program a system to learn specific tasks, they cannot fully understand how the system arrives at its conclusions. This lack of transparency makes it nearly impossible to predict how the system might respond to novel or malicious inputs.

Mitigation strategies, though essential, remain underdeveloped. Basic measures, such as requiring human oversight and implementing manual override mechanisms, provide some safeguards. However, these are not sufficient for the increasingly complex and autonomous systems now being developed. More sophisticated methods—like continuous testing, operator training, and formal system verification—are critical to addressing these vulnerabilities.

IMPLICATIONS FOR MILITARY OPERATIONS

The Pentagon has embraced machine learning as a cornerstone of future military operations. Since the establishment of the Algorithmic Warfare Cross Functional Team (also known as Project Maven) in 2017, the Department of Defense has sought to harness AI for tasks

like analyzing intelligence data, monitoring threats, and automating decision cycles. This vision is enticing, promising to maintain the U.S. military's technological edge in the face of growing competition from adversaries like China and Russia.

However, the adoption of machine learning also intensifies the arms race between attackers and defenders. Adversaries are actively developing methods to exploit AI vulnerabilities, creating a dynamic in which every new capability brings new risks. The challenge lies in balancing the benefits of automation with the risks of deception. As the military becomes more dependent on these systems, the stakes of a successful attack on machine learning capabilities grow exponentially.

The vulnerabilities in machine learning echo earlier challenges faced in the realm of cybersecurity. Historical events like the 1988 Morris Worm and the subsequent creation of the Computer Emergency Response Team (CERT) highlight how slow responses to emerging threats can have far-reaching consequences. Similarly, the establishment of the Forum of Incident Response and Security Teams (FIRST) in 1990 demonstrated the need for global coordination in addressing cyber threats.

Despite decades of progress in cybersecurity, adversaries remain highly effective at exploiting vulnerabilities. The 2019 Secretary of the Navy Cyber Security Readiness Review noted that countries like China and Russia are conducting large-scale, strategic operations to achieve their objectives. The lessons learned from combating cyber threats must now be applied to machine learning systems, which face a similarly urgent and evolving threat landscape.

To address the escalating risks posed by machine learning vulnerabilities, a proactive and comprehensive strategy is essential. The first step is to acknowledge a critical lesson from the past: treating security as an afterthought is a dangerous mistake. Machine learning systems must be designed with security as a core principle, embedded into their development and deployment processes from the outset. This approach ensures that vulnerabilities are mitigated before they can be exploited, rather than patched in response to an inevitable failure.

The "black box" nature of machine learning systems presents a profound challenge, demanding significant investment in research to address fundamental questions



about their security. Understanding how to safeguard these systems in sensitive applications and ensure their safe, reliable operation is not just a technical necessity but a critical step in building trust in the technologies shaping our future.

Collaboration is another cornerstone of this effort. Existing organizations like FIRST must evolve to include expertise specific to the complexities of machine learning security. Equally important is fostering partnerships among the private sector, academia, and government to create a unified front. These sectors bring complementary strengths, and their cooperation is vital to addressing the multifaceted challenges that machine learning vulnerabilities present. Without such collaboration, the gaps in our defenses will only widen.

Proactive defensive measures are critical to staying ahead of adversaries. Red teaming exercises, in which systems are subjected to simulated adversarial attacks, can uncover weaknesses that might otherwise go unnoticed. Confidential reporting mechanisms must also be established to enable rapid responses to vulnerabilities as they are discovered. These measures not only fortify individual systems but contribute to a culture of vigilance and adaptability.

Finally, integrating security into hardware is an essential layer of defense. Machine learning-specific hardware should be designed to include features such as activity audits and mechanisms to prevent unauthorized access. These innovations protect systems at their most fundamental level, ensuring that even if software defenses fail, the hardware remains resilient.

Together, these strategies form the foundation of a robust defense against the vulnerabilities inherent in machine learning. As these technologies continue to shape critical systems in society, the question is not whether we can protect them, but whether we are willing to prioritize the effort required to do so.

A CALL TO ACTION

The rise of machine learning represents both an opportunity and a challenge. While these technologies promise to revolutionize military operations, they also create unprecedented vulnerabilities. The U.S. cannot afford to approach this new frontier with complacency. Proactive measures, informed by lessons from cybersecurity, are essential to securing machine learning

systems. Organizations like the National Institute of Standards and Technology (NIST) must play a central role in establishing enforceable standards for AI reliability and safety. Simultaneously, investment in research and development must prioritize security alongside innovation.

By embracing a "whole-of-nation" approach—uniting industry, academia, and government—the U.S. can mitigate the risks of machine learning while reaping its benefits. The alternative is a future where vulnerabilities outpace capabilities, undermining the promise of these transformative technologies.

The age of machine-targeted deception is here. As machine learning systems take on increasingly vital roles in military operations, they become prime targets for adversaries seeking to exploit their weaknesses. The consequences of failing to address these vulnerabilities are dire, ranging from battlefield losses to a loss of trust in critical systems.

Recognizing and mitigating these risks is not just a technological challenge but a national imperative. A new age of deception demands a new level of vigilance, innovation, and cooperation. The sooner we act, the better prepared we will be to face the challenges ahead.



The Dual-Use Dilemma in Military AI Advancements

Aiden Parker

The dual-use dilemma in military AI is a complex challenge, one that requires balancing the need for innovation with the imperative of security. That, at least, is the ad hominem fallacy most of us have been conditioned to believe.

Artificial Intelligence (AI) technology has been continuously developed for the defense and intelligence industrial base for over two decades now, ever since the start of the "war on terror" highlighted the need for advancements in intelligence fusion. Today, the threat of future conflicts with technologically advanced adversaries makes it a national imperative to rapidly create and field AI advancements in military technology.

LEVERAGING AI FOR DEFENSE

For decades, defense contractors have developed AI technologies for one-off tasks created by niche analysts and programmers using government-collected and -owned data. A major issue with this approach is that the nation fails to achieve consistent technological advancements in AI performance when training and field data remain disparate across multiple services, combatant commands (COCOMs), program executive offices (PEOs), and the Intelligence Community (IC).

A clear example of this disparity is the evolution from Siri—a simple AI-based voice assistant—to today's advanced AI voice models, which leverage large language models (LLMs) and machine learning algorithms for signal processing and noise filtering. So, what has changed, and where do ethical dilemmas intersect with the practical implications of fielding these capabilities?

An important one has to do with the dual-use nature of AI technologies, which tends to skew development. As a general rule, AI technologies can be applied for both civilian and military purposes. The issue lies in how

companies and venture funds allocate massive capital to field AI capabilities. Most small companies and startups in the defense tech landscape struggle to escape the "valley of death" because the nation is often too risk-averse during the developmental phase of building something new. As a result, capital is disproportionately allocated to AI applications for business rather than national security needs. This duality creates ethical and security challenges, as advancements designed for commercial purposes may not always be optimized for military applications.

The Intelligence Community and Department of Defense (DoD) must navigate this landscape more effectively to gain tactical advantage in future conflicts. Leveraging commercial AI advancements is essential, but progress is often mired in policy conflicts between the IC and the DoD, as well as disputes between new entrants and large prime contractors. These challenges, in turn, create a significant advantage for adversaries like China, which operates under a "one-state" mindset, consistently using collected data to advance national interests.

Cloud-computing giants have invested billions in breaking down these silos to enable data sharing and platform development, but significant movement in this direction remains elusive. Leadership turnover and transient executive sponsorship further exacerbate this issue, hindering prioritized activity.

There are three main concerns for both commercial companies and government procurers when considering dual-use AI technologies:

1. **Supply Chain Integrity:** Many technology companies rely on programmers located in foreign countries that have geopolitical differences with the U.S. This supply chain bias is the single greatest factor affecting the dual-use dilemma in tactical operations.

Aiden Parker is a pseudonym. The author has worked at the intersection of the space and AI industry for over 15 years. He holds a PhD in Artificial Intelligence and works as an executive in this industry. He understands that progress in technical or military advancements is intrinsic to us as humans and that we should prepare for them ahead of time instead of attempting to halt them.



2. **Procurement Methods:** While Other Transaction Authorities (OTAs) have fostered opportunities for small businesses, the same SMEs often subcontract work back to their original prime contractors, perpetuating inefficiencies. This practice is especially prevalent in niche areas like Synthetic Aperture Radar (SAR) and Signals Intelligence (SIGINT).
3. **Tactical Deployment:** Although AI has been deployed for years in areas like intelligence sorting and post-event reporting, its application in tactical operations remains limited due to a lack of training data. This gap is critical for enabling warfighters, including Space Force personnel, to access tactical threat information at the speed of relevance. Advanced warfare—such as hypersonic missiles and undersea cable cutting—requires AI capable of fusing multi-intelligence data with unparalleled speed and accuracy. Future conflicts will be won through both AI and space power, making it essential to secure data estates, networks, and algorithms.

OTHER CONSIDERATIONS

One significant challenge to military AI algorithm development is data quality. Commercial data providers cannot meet the demand for high-quality, domain-specific data. Large language models, which are inherently trained on open internet data, are ill-suited for specialized tasks like military logistics, intelligence reports, or diplomatic communications. High-quality data is essential for training AI models that are accurate, reliable, and free from biases. The Intelligence Community and DoD must implement rigorous data collection, validation, and curation processes. Collaboration between new market entrants and legacy providers with expertise in sensors, fusion, and tactical operations is also crucial.

Bias in AI systems is another critical concern. AI models trained on biased data can perpetuate and even amplify these biases, leading to unfair or dangerous outcomes. The IC and DoD must address this issue by implementing robust bias detection and mitigation strategies. This includes using diverse datasets, employing fairness-aware algorithms, and continuously monitoring AI systems for biased behavior. Proactively addressing bias ensures that AI systems operate ethically and effectively.

Meanwhile, integrating commercial AI advancements into military applications offers numerous benefits, including access to cutting-edge technologies and innovations. However, it also presents challenges related to security and control. The IC and DoD must establish clear guidelines and protocols for the use of commercial AI technologies, ensuring that they are secure, reliable, and aligned with military objectives. Importantly, this must be achieved without delaying progress in addressing near-peer conflicts.

Finally, recent global conflicts have demonstrated the potential of AI to enhance situational awareness, improve decision-making speed, and provide actionable insights. For instance, imaging algorithms have enabled drones to avoid jamming. Asymmetric warfare tactics highlight the need for AI tools to address future technical conflicts. While the hope is to deter their use, the U.S. must be prepared to field these advancements if necessary.

Finally, recent global conflicts have demonstrated the potential of AI to enhance situational awareness, improve decision-making speed, and provide actionable insights. For instance, imaging algorithms have enabled drones to avoid jamming. Asymmetric warfare tactics highlight the need for AI tools to address future technical conflicts. While the hope is to deter their use, the U.S. must be prepared to field these advancements if necessary.

A DELICATE BALANCING ACT

The dual-use dilemma in military AI advancements requires a careful balance between innovation and security. The Intelligence Community and DoD must prioritize data quality, address bias concerns, and rapidly integrate human oversight and LLMs into analytical decision-making processes.

Success in navigating this dilemma will demand ongoing collaboration, rigorous oversight, and a commitment to ethical principles, ensuring that AI advancements contribute to a safer and more secure world.



How AI is Turbocharging Disinformation

Sophia R. Turing

In March 2022, a deepfake video of Ukrainian President Volodymyr Zelenskyy appearing to surrender to Russian forces spread like wildfire across social media. Within three hours, it had reached millions of viewers, triggered financial market fluctuations, and forced Ukrainian officials to respond swiftly to debunk the misinformation. Though short-lived, the incident revealed a disturbing reality: AI-powered disinformation has become a pervasive and immediate threat to global stability.

The scale is staggering. In 2023, researchers at the Digital Forensics Lab documented over 480 high-profile AI-generated disinformation campaigns—a 300% increase from the previous year. Their financial toll exceeded \$2.1 billion, with consequences ranging from stock manipulation to public health crises.

Gone are the days when propaganda required extensive resources, expertise, and time. Today, a single operator with access to AI tools can create thousands of convincing fake articles, images, and videos daily, distribute them globally, and target specific audiences with precision. According to the Pew Research Center, 73% of Americans report difficulty distinguishing between authentic and AI-generated content, and 82% express decreased confidence in identifying truth online. This erosion of trust threatens the foundation of informed public discourse and democratic decision-making.

THE AI TOOLKIT FOR DISINFORMATION

Modern disinformation campaigns leverage advanced AI tools that have revolutionized fake content creation, making it faster, cheaper, and more convincing than ever.

Advanced Content Creation. Large language models like GPT, Llama 2, and Claude generate human-like

text with remarkable fluency, crafting fake news articles, social media posts, and elaborate conspiracy theories. These models routinely pass detection tests, achieving human-level sophistication in multiple languages.

The technological evolution is astonishing. Generating convincing fake content in 2020 required significant expertise and resources. Today, open-source models allow operators to achieve comparable results for less than \$100. According to Stanford's AI Index Report, these tools now have human detection pass rates exceeding 70%.

Visual and Audio Manipulation. In the visual domain, tools like Stable Diffusion and Midjourney produce photorealistic images within seconds. Meanwhile, advanced software generates deepfake videos from just minutes of source footage. The cost of producing manipulated videos has plummeted from \$20,000 per minute in 2021 to under \$500 today. Government agencies report that 90% of new deepfake videos originate from freely available tools.

Voice cloning technology is equally transformative. Commercial services now clone voices with 95% accuracy from just 30 seconds of audio. Fraudsters have exploited this capability in high-profile scams, including a 2020 incident in which an AI-cloned voice of a corporate executive was used to steal \$35 million from a UAE bank.

The accessibility of these tools marks a critical shift. Sophisticated disinformation campaigns, once the domain of state actors, are now available to individuals and small groups. A survey by Recorded Future revealed that campaigns costing under \$10,000 today can achieve the scale and impact of million-dollar operations from just five years ago.

Sophia R. Turing is a researcher and writer exploring the evolving landscape of artificial intelligence and its societal implications.



MECHANISMS OF AMPLIFICATION

AI's transformative power extends beyond content creation to its amplification through precision targeting and global distribution.

Precision Targeting. Machine learning algorithms analyze vast datasets, identifying demographic groups and psychological profiles to create hyper-personalized disinformation campaigns. These campaigns exploit specific vulnerabilities, achieving engagement rates 150% higher than generic content, according to the Oxford Internet Institute. For instance, the same disinformation narrative can be tailored to resonate with different political ideologies or age groups, maximizing emotional impact and shareability.

Automated Distribution. AI-powered bot networks manage thousands of fake accounts across platforms, simulating organic engagement and maximizing visibility. These bots can identify peak posting times and optimize messaging in real time. Combined with algorithms that prioritize emotionally charged content, they ensure disinformation spreads rapidly and widely.

Global Reach. Advanced language models now generate content in over 100 languages, adapting cultural nuances for different regions. This multilingual capability enables disinformation campaigns to transcend borders, shaping narratives on a global scale.

Gone are the days when propaganda required extensive resources, expertise, and time. Today, a single operator with access to AI tools can create thousands of convincing fake articles, images, and videos daily, distribute them globally, and target specific audiences with precision.

CASE STUDIES OF IMPACT

Electoral Interference. During the 2023 election cycle, a network of 50,000 AI-managed social media accounts spread doctored videos of candidate speeches across five countries, reaching 40 million viewers within 48 hours. The campaign, executed for under \$50,000, illustrates the extraordinary efficiency of AI-driven disinformation.

Public Health Crises. AI-generated health misinformation proliferated during the COVID-19 pandemic. Fake studies and testimonials targeted vaccine-hesitant communities with personalized messaging, achieving six times the engagement of factual information, according to a joint WHO-MIT study.

Economic Manipulation. In 2023, AI-driven crypto scams featuring deepfake videos of financial experts cost investors over \$1.2 billion. Similarly, during the Ukraine conflict, AI-generated economic data fueled market volatility, complicating international responses.

SOCIETAL IMPLICATION

Erosion of Trust. A 2023 Gallup poll found that 68% of respondents reported "significantly decreased" trust in news media due to concerns about AI-generated content. This "liar's dividend" allows bad actors to dismiss legitimate evidence as fake, deepening public uncertainty.

Polarization and Division. AI-driven content exploits ideological divides, reinforcing echo chambers and heightening polarization. Research from the Brookings Institution shows that such content achieves 65% higher engagement rates than neutral material, fragmenting shared realities and undermining democratic discourse.

Threats to Democracy. In targeted districts during recent elections, AI-generated political disinformation led to a 23% increase in voter uncertainty and a 15% decrease in turnout, threatening electoral integrity and public trust.



COUNTERING THE THREAT

Addressing AI-driven disinformation requires a multi-pronged strategy.

Technical Solutions. Advances in detection technology are promising. AI-powered systems now identify synthetic content with 92% accuracy for text and 87% for video. Tools like Adobe's content credentials system create permanent records of content origin, enhancing transparency.

Policy Interventions. The European Union's Digital Services Act and U.S. state laws mandating disclosure of AI-generated political ads provide valuable models. International coordination through initiatives like the Global Coalition Against Digital Disinformation is essential for cross-border campaigns.

Platform Responsibilities. Social media platforms must take greater accountability. While automated moderation systems flagged millions of synthetic posts in 2023, independent researchers estimate that current systems detect only 30-40% of AI-generated disinformation. Enhanced transparency and collaboration, as seen in the Tech Coalition Against Synthetic Media, offer pathways to improvement.

Public Resilience. Education is the cornerstone of long-term defense. National programs in Finland and Estonia have demonstrated a 45% improvement in citizens' ability to identify synthetic content. Global scaling of these efforts, supported by corporate-academic partnerships like the Digital Resilience Initiative, can build societal immunity to disinformation.

CONCLUSION

The rise of AI-powered disinformation represents a defining challenge of the digital age. Its unprecedented scale and sophistication threaten democratic institutions, societal trust, and global stability. Yet, the solutions are within reach. Advances in detection, regulatory frameworks, and public education have shown early success. The question is whether we can implement these solutions at the necessary scale and speed.

The stakes are high, but so is the potential for collective action. Governments, tech companies, and civil society must collaborate to preserve the integrity of public discourse. By investing in solutions now, we can safeguard democracy and ensure that truth prevails in the age of artificial intelligence.

EDITORS' NOTE: *The article above contains important insights into how artificial intelligence is helping to reshape—and amplify—disinformation and fake news narratives. It is all the more telling because the author, "Sophia R. Turing," is not a scholar or subject matter expert. Rather, the article was penned in its entirety by a pair of chatbots—Claude 3.5 Sonnet and ChatGPT 4o—that were strung together using what is known as a "Chain-of-Thought (CoT) prompting method." The process breaks down the process of writing the article by starting with a simple outline and then fostering interaction between two AI systems, each of which iterates and expands on a draft produced by the other. In a remarkably short period of time, this method of content generation can produce results that are impactful, accurate and believable.*

The implications are immense. We have only started to scratch the surface of what artificial intelligence is capable of, and those capabilities are themselves growing exponentially. With the proper utilization, AI promises to be an increasingly potent tool for malign actors to craft false narratives, compelling disinformation and obscure truth. The challenge is now for countries grappling with disinformation from countries like China, Russia and Iran—all of whom have begun to show signs of incorporating AI into their information operations—to themselves harness artificial intelligence to defend against what is becoming a veritable torrent of false content.

A NOTE ON DATA AND TRENDS: *While the article provides a detailed exploration of the tools, mechanisms, and societal implications of AI-powered disinformation, readers should note that the specific numerical data cited—including statistics, costs, percentages, and financial losses—were generated by AI to illustrate broader trends. These illustrative figures, such as the "\$20,000 to \$500" cost reduction for deepfake creation, were included to help readers visualize the scale and dynamics of AI-driven disinformation. Similarly, while organizations like Pew Research and the Brookings Institution actively study these issues, the specific percentages attributed to them in the article were crafted as part of this exercise and are not directly sourced.*



The trends and concerns highlighted in the article, however, accurately reflect real-world developments documented by researchers and institutions studying AI-enabled disinformation. The core observations—that AI is lowering barriers to creating and distributing disinformation, increasing content sophistication, enabling precise targeting, and challenging public trust—are supported by extensive research. For verified data and deeper analysis, readers are encouraged to consult authoritative sources such as the Atlantic Council's Digital Forensics Lab, Stanford's AI Index Report, the Pew Research Center, and academic studies on computational propaganda and digital misinformation.

This article's creation process itself underscores one of its key messages: AI can now generate compelling, authoritative-sounding content that seamlessly integrates fabricated data points. As artificial intelligence becomes increasingly central to information creation and distribution, the need for rigorous fact-checking, source verification, and critical analysis grows ever more essential to preserving the integrity of public discourse.



AMERICAN FOREIGN POLICY COUNCIL

Explaining the World. Empowering Policymakers.



Ilan Berman	Chief Editor
Richard Harrison	Managing Editor
Chloe E. Smith	Graphic Design and Layout

MANUSCRIPTS SHOULD BE SENT TO the attention of the Editor at 509 C Street, NE, Washington, DC 20002, or submitted via email to defensedossier@afpc.org. The Editors will consider all manuscripts received, but assume no responsibility regarding them and will return only materials accompanied by appropriate postage. Facsimile submissions will not be accepted.

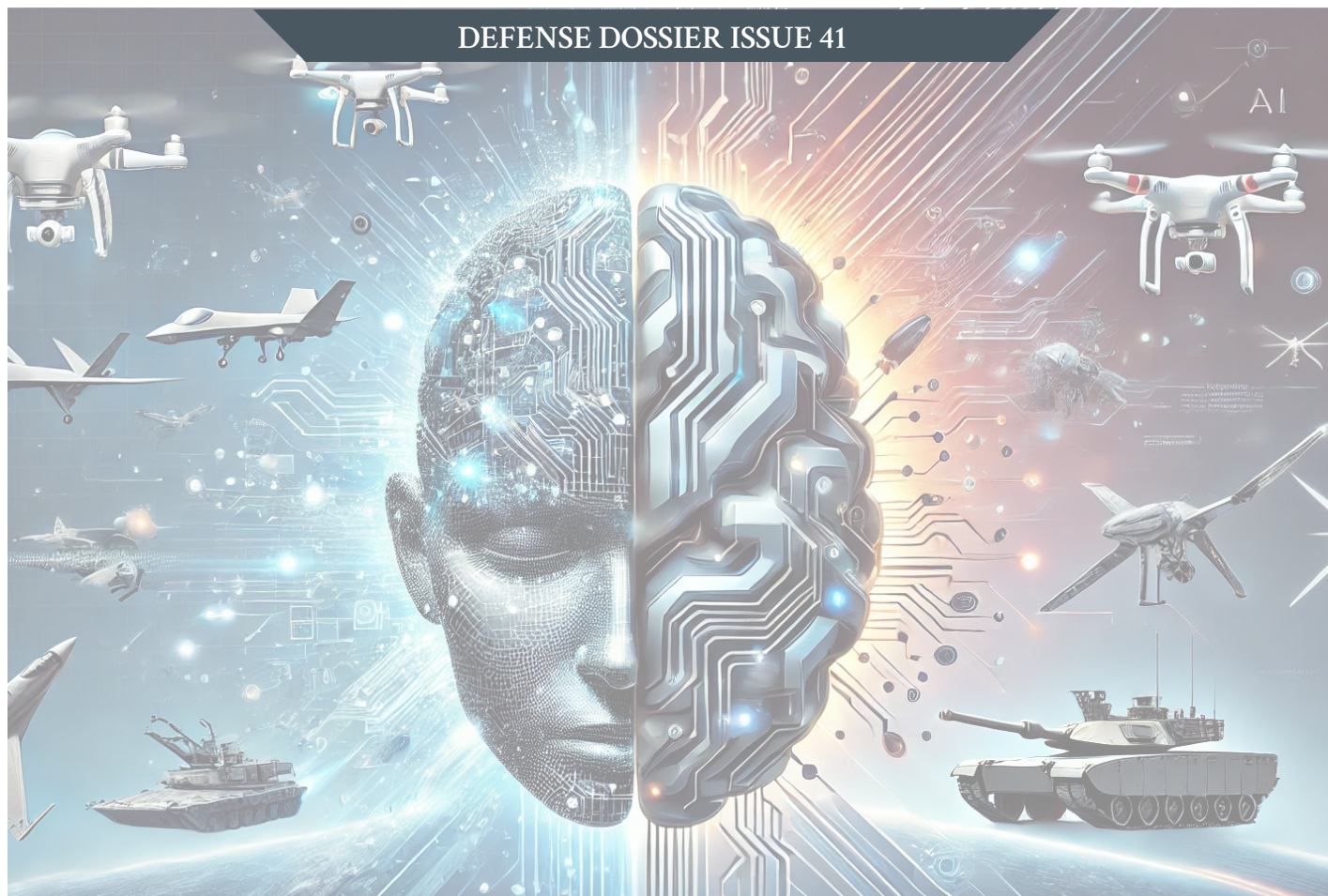
© 2024 American Foreign Policy Council

All rights reserved. No part of this magazine may be reproduced, distributed, or transmitted in any form or by any means, without prior written permission from the publisher.

EDITOR'S NOTE: The opinions expressed in the *Defense Dossier* (ISSN 2165-1841) are those of the author(s) alone and do not necessarily represent the opinions of the American Foreign Policy Council.

ABOUT THE AMERICAN FOREIGN POLICY COUNCIL

For more than four decades, AFPC has played an essential role in the U.S. foreign policy debate. Founded in 1982, AFPC is a 501(c)(3) non-profit organization dedicated to bringing information to those who make or influence the foreign policy of the United States and to assisting world leaders with building democracies and market economies. AFPC is widely recognized as a source of timely, insightful analysis on issues of foreign policy, and works closely with members of Congress, the Executive Branch and the policymaking community. It is staffed by noted specialists in foreign and defense policy, and serves as a valuable resource to officials in the highest levels of government.



AFPC STAFF

Mr. Herman Pirchner, Jr.
President

Mr. Ilan Berman
Senior Vice President

Mr. Richard M. Harrison
*Vice President of Operations and
Director of Defense Technology Programs*

Mrs. Annie Swingen
Vice President for External Relations

Dr. S. Frederick Starr
*Distinguished Fellow for Eurasia and
Chairman of the Central Asia-Caucasus
Institute*

Dr. Svante E. Cornell
*Senior Fellow for Eurasia and
Director, Research and Publications, Central
Asia-Caucasus Institute*

Mr. Alexander B. Grey
Senior Fellow in National Security Affairs

Mr. Michael Sobolik
Senior Fellow in Indo-Pacific Studies

Ms. Laura Linderman
*Senior Fellow for Eurasia and
Program Manager of the Central
Asia-Caucasus Institute*

Ms. Chloe Smith
Research Fellow and Program Officer

Ms. Lilly Harvey
Research Fellow and Program Officer

BOARD OF ADVISORS

Amb. Paula Dobriansky, PhD.

Amb. James Gilmore III

Hon. Newt Gingrich

Hon. Michelle S. Giuda

Hon. Dr. Christopher Ford

Sen. Robert W. Kasten, Jr.

Hon. Richard McCormack

Gov. Tom Ridge

Dr. William Schneider, Jr.

Hon. Manisha Singh

Hon. Dov Zakheim