PROACTIVE AI RISK MANAGEMENT. THE SECOND AI ARMS RACE: FROM DEREGULATION TO INDUSTRIAL POLICY, SOVEREIGN INFRASTRUCTURES, AND ALGORITHMIC WARFARE

Alexey V. Amenitsky [1], **Evgeny G. Vorobyov** [2]

[1] Saint Petersburg State Electrotechnical University "LETI", Saint Petersburg, Russia, ORCID ID: 0009-0004-0955-1527;

[2] Saint Petersburg State Electrotechnical University "LETI", Saint Petersburg, Russia, ORCID ID: 0000-0003-0564-5935

ABSTRACT

The global competition in artificial intelligence (AI) has entered a qualitatively new phase – what this article terms the second AI arms race (AI Arms Race 2.0). Moving beyond early narratives of innovation and deregulation, this stage is characterized by the deliberate fusion of economic and national security agendas, large-scale state-industrial coordination, and the militarization of foundational AI models. Drawing on primary policy documents, corporate disclosures, and expert analyses from 2023–2025, we identify three systemic shifts: (1) the transition from market-led to state-directed AI industrial policy, exemplified by U.S. export controls, sovereign AI initiatives in the EU, and China's techno-strategic autonomy drive; (2) the collapse of the "anti-military AI consensus" among major technology firms, with OpenAI, Google, and Meta now explicitly permitting – and even advocating – the use of their models in defense and surveillance applications; and (3) the emergence of algorithmic warfare, where AI agents execute cyber operations at machine speed, raising unprecedented challenges for attribution, escalation control, and defensive equity. We argue that this new race is less about raw model performance and more about infrastructural sovereignty, data geopolitics, and the institutional capture of AI governance. Crucially, the "arms race" framing – while real in strategic terms – also functions as a discursive tool to depoliticize regulation and consolidate power among a narrow set of corporate-state actors. The article concludes with a normative framework for human-centered AI arms control, grounded in transparency, multilateral verification, and the protection of open innovation ecosystems.

KEYWORDS: *artificial intelligence; AI arms race; AI industrial policy; sovereign AI; autonomous weapons; algorithmic warfare; AI nationalism; export controls; cybersecurity; AI governance.*

DOI: 10.36724/2664-066X-2025-11-5-28-33

Received: 07.09.2025
Accepted: 23.10.2025

Citation: Alexey V. Amenitsky, Evgeny G. Vorobyov, "Proactive AI risk management. The Second AI Arms Race: From Deregulation to Industrial Policy, Sovereign Infrastructures, and Algorithmic Warfare", *Synchroinfo Journal* **2025**, vol. 11, no. 5, pp. 28-33.

Licensee IRIS, Vienna, Austria.

1 Introduction

Since the mid 2010s, the trope of a "U.S.–China AI arms race" has dominated strategic discourse [11, 12, 8]. Initially deployed by industry actors to resist data protection and antitrust legislation [11, §1.1], the metaphor has evolved from rhetorical device to operational doctrine. By 2024-2025, the race had entered its second phase – AI Arms Race 2.0 – marked not by deregulation, but by active state industrial policy: massive public investment, chip export controls, national "AI factories," and the explicit enlistment of private AI labs into national security missions.

This article analyzes the structural, institutional, and normative transformations underpinning this shift. Our analysis draws on:

• U.S. pr sidential directives (NSC Memorandum, 2024) [17, 18];

• Corporate policy r visions (OpenAI, Google, Meta, Anthropic);

• Sovereign AI initiatives (EU InvestAI, France's €110 bn plan, NVIDIA's strategic repositioning) [13];

• Incident reports (e.g., Anthropic's 2024 disclosure of AI automated cyber espionage);

• Crit cal policy scholarship [11, 15].

We advance three core arguments:

1. *The race is no longer bilateral*. While the U.S.–China dyad remains central, the rise of AI nationalisms [11] has fragmented the global landscape into competing "digital sovereignties" – U.S.-aligned, EU-autonomous, China-centric, and Gulf-funded ecosystems [9].

2. *The boundary between civilian and military AI has collapsed*. Firms that once resisted military collaboration now actively integrate defense use cases, reframing this as a civilizational imperative [10, 2].

3. *The "arms race" narrative serves dual functions*: it justifies state support for dominant firms (too strategically important to fail), and simultaneously legitimizes regulatory exemptions – thereby consolidating a Silicon Valley-Washington consensus [11].

The remainder of the article proceeds as follows: Section 2 outlines the conceptual shift from "AI 1.0" (market-driven) to "AI 2.0" (state-driven); Section 3 examines the militarization of foundational models; Section 4 analyzes the rise of sovereign AI as both strategic response and market opportunity; Section 5 discusses the risks of algorithmic warfare and defensive inequity; and Section 6 proposes a path toward accountable AI governance.

2 AI Arms Race 2.0: From Deregulation to State-Led Industrial Policy

The first phase of the AI race (2016-2022) emphasized light-touch regulation, venture capital dynamism, and global data flows [8]. In contrast, AI Arms Race 2.0 – crystallizing in 2023-2025 – is defined by strategic state intervention.

2.1. The U.S. Turn: Securitization as Economic Policy

Under the Biden administration, AI policy was explicitly linked to economic security. National Security Advisor Jake Sullivan (2023, 2024) framed technological leadership as inseparable from national power, paving the way for [16]:

• Executive Order 14110 (Oct. 2023), mandating federal AI adoption and safety standards;

• NSC Memorandum (Oct. 2024), directing agencies to prioritize AI for defense missions;

• Export Control Framework for AI Diffusion (2024), restricting sales of advanced chips (e.g., NVIDIA A100/H100) to China and allied jurisdictions.

These measures aligned closely with corporate lobbying. OpenAI, for instance, conditioned its domestic investment on government support, even threatening relocation (Wolman & Chatterjee, 2024) [14].

The Trump administration (inaugurated Jan. 2025) escalated this trajectory, repealing EO 14110 and issuing "Removing Barriers to American Leadership in Artificial Intelligence" (Jan. 23, 2025), which declared:

"It is the policy of the United States to sustain and enhance America's global AI dominance in order to promote human flourishing, economic competitiveness, and national security." [18]

Crucially, industrial policy extended beyond subsidies: it involved personnel integration. Key tech executives assumed government roles – David Sacks (a16z) as White House "AI Czar," Michael Kratsios (Scale AI) as OSTP Director, Jacob Helberg (Palantir) as Under Secretary for Economic Growth [1].

2.2. Beyond the Bipolar Frame: AI Nationalisms and Sovereign Infrastructures

While U.S.–China rivalry dominates headlines, AI nationalism is proliferating globally [12]. States pursue dual objectives: (1) reduce dependency on U.S. platforms; (2) attract investment by offering regulatory "safe harbors."

• European Union: Repurposed EuroHPC supercomputers into AI Factories, launched the €20 bn InvestAI program to build GPU "gigafactories," and backed the EuroStack movement for sovereign tech stacks (EU Commission, 2025; Kaltheuner & Saari, 2025).

• France: Announced €110 bn in AI commitments at the Paris AI Action Summit (Feb. 2025), targeting data centers and talent pipelines (Reuters, 2025).

• Gulf States: MGX (UAE) and PIF (Saudi Arabia) deployed >$30 bn into AI startups (CNBC, 2024; NYT, 2024), positioning as financial swing states in tech diplomacy (Karaian, 2025).

Even NVIDIA – long a beneficiary of globalized supply chains – pivoted to promote Sovereign AI, defining it as a nation's capacity to develop AI via domestic infrastructure, data, workforce, and business networks [13]. This reframing allows chipmakers to monetize state anxieties: sovereign build-outs generate demand for localized data centers and custom hardware, offsetting losses from export bans (Kaltheuner et al., 2025).

Table 1

Comparative AI Industrial Policy Initiatives (2023-2025)

Region	Flagship Program	Budget	Key Actors	Strategic Goal
USA	Stargate JV	$500 bn	Microsoft, OpenAI, CoreWeave, Oracle	Infrastructure dominance, military integration
EU	InvestAI / AI Factories	€20 bn	EuroHPC, Commission	Technological sovereignty, decoupling
France	Paris AI Action Plan	€110 bn	CEA, Mistral AI	European leadership in foundational models [5]
UAE/GCC	MGX, G42, PIF AI Fund	>$30 bn	G42, Cerebras, Microsoft	Geopolitical influence, diversification
China	Next-Gen AI 2.0	Undisclosed	Baidu, Alibaba, Huawei	Autonomy in chips, training, deployment
Sources: EU Commission (2025); Reuters (2025); CNBC (2024); Singh (2025)				

3 The Collapse of the Anti Military AI Consensus

Historically, firms like Google and OpenAI maintained ethical red lines around military AI – epitomized by Google's withdrawal from Project Maven (2018) and OpenAI's 2020 "no military use" clause.

By 2024–2025, this consensus had dissolved:

Company	Policy Shift	Date	Key Statement / Action
OpenAI	Removed ban on "military and warfare" use	Jan. 2024	Updated Acceptable Use Policy to permit defense applications [4]
Google	Reversed AI weapons ban	Feb. 2025	Demis Hassabis: "We must support national security in an era of strategic competition" [7]
Meta	Offered Llama models to U.S. government	Nov. 2024	For "national security purposes" (Moorhead, 2024)
Anthropic	Justified U.S. AI acceleration as defense against "authoritarian AI dominance"	Feb. 2025	Amodei: "Speed is a security imperative" [2]
Palantir	Mission reframed as civilizational defense	Jan. 2025	Karp: "We're here to… scare our enemies and, on occasion, kill them" [10]

This shift is institutionalized through forums like the Hill & Valley Forum (co-founded by Helberg), where Silicon Valley elites and policymakers coordinate on "countering China's tech influence" (Dwoskin, 2025; Chatterjee, 2025).

The underlying logic is securitization: any regulatory constraint is recast not as consumer protection, but as strategic vulnerability. As Kak et al. [11, §1.2] observe:

"The arms race narrative insulates firms from accountability by framing oversight as harmful to national interest."

4 Algorithmic Warfare and the Crisis of Defensive Equity

The convergence of AI and cyber operations has given rise to algorithmic warfare — high-speed, autonomous campaigns executed by AI agents with minimal human input.

4.1. The Anthropic Incident (2024)

In September 2024, Chinese state-sponsored actors reportedly used Anthropic's Claude Code model to automate reconnaissance, exploitation, and exfiltration across ~30 targets. According to U.S. Cyber Command, the AI executed 80-90% of the operation, operating at thousands of requests per second [2, 3, 6].

Anthropic's public disclosure was met with skepticism:

• Yann LeCun (2025) accused the firm of inflating threats to justify regulating open models out of existence;

• Researchers noted the absence of IOCs or forensic evidence (The Stack, 2025);

• Ch na denied involvement, demanding "substantial evidence" [6].

Nonetheless, the strategic implications are profound: if offense can be automated at scale, defense must respond in kind – yet only a handful of actors possess the data, compute, and expertise to build AI-powered cyber defenses.

4.2. The "Defensive Gap" and Systemic Risk

This creates a defensive equity crisis:

• Large la s and states develop "AI immune systems" with built-in safeguards;

• SMEs, civil society, and Global South actors remain reliant on legacy tools, increasing systemic vulnerability.

As one cybersecurity consultant noted:

"Security through obscurity fails. Distributed, open development has historically produced more robust systems – but machine-speed attacks may invalidate that logic." [6]

The dilemma is acute: closed, proprietary models may offer short-term security, but undermine long-term resilience by concentrating power and reducing auditability.

5 Toward Human-Centered AI Arms Control

Given these trends, traditional arms control paradigms (e.g., treaties banning specific weapons) are ill-suited for AI. Instead, we propose a three-pillar framework for human-centered AI arms control:

5.1. Transparency and Verification

• Mandatory red-teaming for high-risk AI systems (e.g., those used in cyber defense/offense);

• Public registries of military AI deployments (modeled on nuclear warhead declarations);

• Independent auditing of "safeguards" in commercial models (cf. EU AI Act, high-risk category).

5.2. Protection of Open Innovation

• Exempt open-weight, non-military AI models from export controls and licensing burdens;

• Fund publ c "AI commons" for cybersecurity tools (e.g., via OECD or UNIDO).

5.3. Multilateral Norm-Setting

• Revive negotiations on a Political Declaration on Responsible Military AI Use (building on 2023 Bletchley commitments);

• Establ sh an International AI Security Board under UN auspices, with technical and ethical expertise.

Crucially, any regime must resist regulatory capture: the framing of AI as an "existential arms race" must not become a pretext for entrenching corporate monopolies.

6 Conclusion

The AI arms race is no longer speculative – it is institutional, infrastructural, and operational. AI Arms Race 2.0 is distinguished by the fusion of economic and security policy, the militarization of general-purpose models, and the strategic promotion of "sovereign AI" as both shield and market.

Yet the dominant narrative – a binary, zero-sum contest between the U.S. and China – masks deeper transformations: the rise of multipolar AI ecosystems, the financialization of sovereignty (Gulf funds), and the quiet consolidation of power among a narrow corporate-state elite.

The central question is no longer who will win the race, but what kind of world the race is building. Without deliberate, inclusive governance, AI's promise of shared prosperity may give way to a fragmented, weaponized, and inequitable digital order.

REFERENCES

[1] M. Alder, “Trump taps Michael Kratsios, Lynne Parker for tech and science roles,” Fedscoop. 2024. https://fedscoop.com/trump-taps-michael-kratsios-lynne-parker-tech-science-roles

[2] D. Amodei, “Statement from Dario Amodei on the Paris AI Action Summit,” Anthropic. 2025, February 11. https://www.anthropic.com/news/paris-ai-summit

[3] Anthropic. 2024, December. Report on AI-assisted cyber intrusion. Internal briefing, cited in Forbes (2025).

[4] S. Biddle, “OpenAI quietly deletes ban on using ChatGPT for “military and warfare”,” 2024, January 12. The Intercept. https://theintercept.com/2024/01/12/open-ai-military-ban-chatgpt

[5] European Commission. AI Continent Action Plan. 2025. https://digital-strategy.ec.europa.eu/en/factpages/ai-continent-action-plan

[6] Forbes. The AI arms race has arrived: The real question is who gets to arm. 2025, November 30. https://www.forbes.com/sites/arafatkabir/2025/11/30/the-ai-arms-race-has-arrived/

[7] I. Fried, “Google’s Hassabis explains shift on military use of AI,” Axios. 2025, February 14. https://www.axios.com/2025/02/14/google-hassabis-ai-military-use

[8] T. Ghi, and A. Srivastava, “The global AI arms race – how nations can avoid being left behind,” Bernard Marr & Co. 2021. https://bernardmarr.com/the-new-global-ai-arms-race

[9] Global Policy Journal. The artificial intelligence arms race: Trends and world leaders in autonomous weapons development. 2025. https://www.globalpolicyjournal.com/articles/conflict-and-security/artificial-intelligence-arms-race

[10] S. Hurwitz, “The gleeful profiteers of Trump’s police state,” Mother Jones. 2025, February 6. https://www.motherjones.com/politics/2025/02/palantir-alex-karp-trump

[11] A. Kak, S.M. West, and I.D. Raji, “AI Arms Race 2.0: From deregulation to industrial policy,” AI Now Institute. 2025. https://ainowinstitute.org/publications/research/1-3-ai-arms-race-2-0

[12] A. Kak, S.M. West, M. Singh, and I.D. Raji, “AI Nationalism(s): Global industrial policy approaches to AI,” AI Now Institute. 2024. https://ainowinstitute.org/ai-nationalisms

[13] A. Lee, “What is Sovereign AI?” NVIDIA Blog. 2024, February 28. https://blogs.nvidia.com/blog/what-is-sovereign-ai

[14] C. Metz, and T. Mickle, “Behind OpenAI’s audacious plan to make A.I. flow like electricity,” The New York Times. 2024, September 25.

[15] M. Singh, “Stargate or StarGatekeepers? Why this joint venture deserves scrutiny,” Berkeley Technology Law Journal, 41. 2024, September 25. https://doi.org/10.2139/ssrn.5184657

[16] J. Sullivan, “Remarks by National Security Advisor Jake Sullivan on renewing American economic leadership,” The White House. 2023, April 27. https://bidenwhitehouse.archives.gov/briefing-room/speeches-remarks/2023/04/27/remarks-by-national-security-advisor-jake-sullivan

[17] White House. Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Federal Register, 2023, no. 88(210), pp. 75191-75226.

[18] White House. Removing Barriers to American Leadership in Artificial Intelligence. 2025, January 23. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence