Forged in Fire

How Artificial Constraints Breed Technological Dominance and Why the Only Winning Strategy in the AI Race Is Safety

Travis Gilly
Real Safety AI Foundation
t.gilly@ai-literacy-labs.org

White Paper | February 2026

Abstract

This white paper examines a consistent historical pattern in which powerful entities impose artificial constraints on rising competitors, only to accelerate the innovation they sought to prevent. Drawing on case studies from military conflict, industrial competition, economic sanctions, and contemporary AI policy, we demonstrate that constraint-driven efficiency gains persist as permanent competitive advantages even after the original constraints are removed. We apply this pattern to the ongoing U.S.-China AI competition, where semiconductor export controls have demonstrably accelerated Chinese AI innovation rather than impeding it, as documented by Brookings, RAND, CSIS, and other major policy institutions. We then present a three-branch strategic analysis demonstrating that regardless of China's intentions (cooperative, self-destructive, or adversarial), the optimal U.S. strategy converges on the same conclusion: building AI systems with rigorous safety engineering. We argue that safety constraints function as a voluntary version of the same mechanism that produces constraint-driven innovation, yielding systems that are not weaker for their discipline but more resilient, more predictable, and more defensible.

Keywords

AI safety; export controls; semiconductor policy; constraint-driven innovation; DeepSeek; U.S.-China AI competition; interpretability; safety engineering; national security; Harm Blindness Framework; lean manufacturing; guerrilla innovation; distributed systems; AI governance; technology policy

1. Introduction

In October 2022, the United States Bureau of Industry and Security issued sweeping export controls on advanced semiconductor technology to the People's Republic of China. The stated objective was to deny Chinese AI developers access to the computational infrastructure required to train frontier models, thereby preserving American technological leadership in artificial intelligence. Within three years, every major policy institution that studied the results arrived at a version of the same conclusion: the controls were not working as intended.

Brookings found that the restrictions may actually be accelerating Chinese AI innovation.[1] CSIS documented genuine technological breakthroughs emerging from Chinese labs operating under constraint.[2] RAND noted that Huawei's domestically produced CloudMatrix system outperforms Nvidia's state-of-the-art GB200 system on key performance dimensions.[4] DeepSeek, a Chinese startup, trained models competitive with GPT-4 at a fraction of the cost, using chips whose performance had been deliberately capped by export regulations.[11, 12]

This outcome should not have surprised anyone. The pattern of artificial constraints breeding accelerated innovation is not novel; it is among the most thoroughly documented dynamics in military, economic, and industrial history. From guerrilla warfare to automotive manufacturing, from pharmaceutical development under embargo to semiconductor supply chain restructuring, the historical record is unambiguous. Entities that are forced to do more with less develop engineering cultures, institutional knowledge, and optimization reflexes that persist as permanent competitive advantages long after the original constraints are removed.

This white paper traces that pattern across five historical case studies, maps it onto the current U.S.-China AI dynamic, and then asks the question that the policy literature has largely avoided: if constraints breed better engineering, what does that imply for AI safety? We argue that the answer is both counterintuitive and strategically decisive. Safety engineering is not the drag on innovation that its critics claim; it is a chosen constraint that produces the same efficiency and resilience advantages that forced constraints produce. The difference is that chosen constraints come without the geopolitical risk of teaching your competitor to thrive on adversity.

2. The Theory of Constraint-Driven Innovation

The observation that necessity drives invention is ancient, but the mechanism deserves more precise articulation than the proverb provides. Constraint-driven innovation operates through three distinct phases. In the first phase, an entity facing resource limitations redirects engineering effort from brute-force solutions toward efficiency optimization. In the second phase, the accumulated efficiency gains become embedded in institutional culture, training pipelines, and engineering doctrine. In the third phase, when constraints are removed or relaxed, the entity applies its efficiency-optimized approach to newly available resources, producing outcomes that exceed what either unconstrained brute force or constrained efficiency could achieve alone.

The critical insight is in the third phase. The efficiency culture does not disappear when the constraint does. A team that learned to extract maximum performance from limited hardware does not forget how to do so when better hardware arrives. Instead, they apply the same optimization discipline to superior resources, compounding both advantages simultaneously.

This dynamic is well-documented across domains. In engineering, in military doctrine, in manufacturing, and now in AI development, the pattern holds: teams forced to optimize do not stop optimizing when the constraint lifts. They apply their hard-won efficiency to better resources, compounding both advantages simultaneously. The question this paper poses is whether that same dynamic applies when the constraint is chosen rather than imposed.

3. Historical Evidence

3.1 Japanese Automotive Manufacturing (1950s-1980s)

Post-war Japan faced acute resource scarcity: limited raw materials, constrained factory space, and insufficient capital for the mass-production approaches that defined American automotive dominance. Toyota's response was not to build a worse version of a Ford factory; it was to invent an entirely different manufacturing philosophy. The Toyota Production System, encompassing just-in-time inventory, continuous improvement (kaizen), and elimination of waste (muda), emerged directly from the constraint of having to produce competitive vehicles without American-scale resources.[17]

When the 1973 oil crisis shifted consumer demand toward fuel-efficient vehicles, Japanese manufacturers did not merely happen to have the right product. They had spent two decades building an engineering culture optimized for efficiency at every level of production. American manufacturers, accustomed to abundant resources and large-margin vehicles, could not replicate that culture overnight. The constraint-bred advantage persisted for decades. By the 1980s, Japanese automotive quality and efficiency had become the global benchmark, and the lean manufacturing principles born from scarcity became the dominant paradigm worldwide.

The lesson is not simply that Japan made good cars. The lesson is that the efficiency discipline, forged under constraint, became a permanent institutional advantage that compounded when resources became available. Japan did not revert to wasteful production when it could afford to. The constraint had changed the culture.

3.2 Vietnam and Guerrilla Warfare (1955-1975)

The Viet Cong and North Vietnamese Army operated with dramatically inferior hardware against the most technologically advanced military in the world. American forces deployed B-52 bombers, napalm, Agent Orange, and a logistical infrastructure that dwarfed anything available to the Vietnamese. The constraint was existential, not merely competitive.

The response was not to build inferior versions of American weapons. It was to develop an entirely different doctrine. Tunnel networks extending hundreds of kilometers, supply lines concealed beneath

jungle canopy, logistical systems that moved materiel by bicycle and foot, and tactical approaches that negated American technological superiority through proximity, patience, and terrain exploitation. Every resource was maximized because no resource could be wasted.

Viewed through an engineering lens, the Vietnamese built a resilient distributed network to counter a centralized high-throughput system. American power depended on concentrated assets: air bases, carrier groups, B-52 sorties launched from fixed positions with massive logistical footprints. Vietnamese power was decentralized and lightweight: the Ho Chi Minh trail functioned as a fault-tolerant supply network where any single node could be destroyed without collapsing the system. The parallel to contemporary computing architectures is not metaphorical; it is structural. Centralized brute-force compute (massive GPU clusters running unoptimized code) is vulnerable in exactly the ways that centralized military power was vulnerable: expensive to maintain, catastrophic when disrupted, and fundamentally less adaptive than distributed alternatives.

When Soviet and Chinese military hardware did flow into Vietnam in greater quantities, the Vietnamese did not abandon their hard-won tactical efficiency. They combined guerrilla doctrine with conventional capability, producing a force that could operate across the full spectrum of conflict. The constraint had not merely produced workarounds; it had produced a military culture that extracted maximum strategic value from every available asset. That culture persisted and compounded when assets increased.

3.3 Israel and the Defense Industry (1948-Present)

Israel's early decades were defined by arms embargoes from multiple nations, including France (which had been a primary supplier) following the 1967 Six-Day War. Facing existential security threats without reliable access to external arms markets, Israel invested heavily in domestic defense capabilities. The constraint was not optional; it was survival.[18]

The result was not merely self-sufficiency but global leadership. Israel developed the Iron Dome missile defense system, the Merkava tank, the Lavi fighter jet program (later contributing to the F-35), and one of the most sophisticated signals intelligence capabilities in the world. The defense technology sector became the foundation for a broader technology ecosystem, producing the highest density of startups per capita of any nation. The efficiency and innovation culture born from constrained defense needs extended into civilian technology, cybersecurity, and artificial intelligence.

Israel today is one of the world's leading arms exporters. The constraints that forced domestic innovation did not produce an inferior defense industry; they produced a superior one, because every system had to work within tighter tolerances, every design had to account for limited production runs, and every engineering decision carried existential weight.

3.4 Cuba and Pharmaceutical Innovation Under Embargo (1960-Present)

The U.S. trade embargo on Cuba, beginning in 1960, cut the island nation off from pharmaceutical imports and medical technology. Rather than simply experiencing degraded healthcare, Cuba invested

heavily in domestic biotech research and development, producing results that embarrassed nations with vastly greater resources.[19]

Cuba developed CIMAvax, a lung cancer vaccine, under conditions that no Western pharmaceutical company would tolerate. The nation produces and exports vaccines to dozens of countries across Latin America and Africa. Cuban medical training, necessitated by the inability to import sufficient healthcare professionals, created one of the highest doctor-to-patient ratios in the world and a medical diplomacy program that has deployed physicians to over 60 countries.

The pharmaceutical innovation was not incidental to the constraint; it was caused by it. When you cannot import the medicine, you learn to make it. When you learn to make it under severe resource limitations, you develop manufacturing processes that are efficient enough to export. The constraint did not produce inferior medicine; it produced medicine optimized for resource-limited environments, which happens to be what most of the world actually needs.

3.5 South Korea and Semiconductor Supply Chains (2019-Present)

In July 2019, Japan restricted exports of three key semiconductor materials to South Korea amid a diplomatic dispute. The restrictions targeted fluorinated polyimide, photoresists, and hydrogen fluoride, all critical to semiconductor manufacturing. South Korean chipmakers Samsung and SK Hynix, initially dependent on Japanese suppliers for these materials, faced an immediate supply chain vulnerability.[20]

Within two years, South Korean companies had dramatically reduced their dependence on Japanese materials, developing domestic suppliers and alternative sourcing relationships. The constraint forced a supply chain restructuring that, once completed, left Samsung and SK Hynix more resilient than before the restrictions were imposed. Japan's leverage evaporated precisely because it was exercised. The attempt to create dependency instead created independence.

This case is particularly instructive because the timeline is compressed and well-documented. The constraint was imposed, the innovation response occurred, and the constraint became irrelevant, all within roughly 24 months. The speed of adaptation suggests that the constraint-to-innovation pipeline operates faster in technology-intensive industries than historical precedents might suggest.

4. The Current Case: China, Chip Bans, and the DeepSeek Phenomenon

The historical pattern is now repeating in artificial intelligence with remarkable fidelity. U.S. export controls have restricted China's access to advanced Nvidia GPUs, the hardware that powers frontier AI model training. The intended effect was to slow Chinese AI development. The actual effect, documented across multiple independent analyses, has been to accelerate Chinese innovation in AI efficiency.

DeepSeek, a Chinese AI startup, trained its V3 foundation model using Nvidia H800 chips (performance-capped versions designed to comply with export regulations) at a reported cost of approximately $6 million, compared to over $100 million for OpenAI's GPT-4.[13] The company's engineers, unable to access H100 chips, developed novel approaches to reduce memory usage and accelerate computation without sacrificing accuracy. As a former DeepSeek employee described it, the team's culture centered on turning hardware challenges into innovation opportunities.[12]

The efficiency innovations were not superficial. DeepSeek programmed 20 of the 132 processing units on each H800 specifically to manage cross-chip communications, working at a lower programming level than Nvidia's standard CUDA platform.[2] This is precisely the kind of deep optimization that emerges from constraint; it requires understanding the hardware at a level that well-resourced teams rarely bother to explore because they do not need to.

Brookings stated the dynamic plainly: scarcity fosters innovation, and when China eventually develops the ability to produce its own leading-edge computing chips, it will possess both computing capacity and efficient algorithms.[1] CSIS acknowledged that DeepSeek demonstrated genuine breakthroughs at or near the absolute frontier of foundational AI research, accomplishments achieved by a team educated and trained almost entirely within China.[2]

The hardware trajectory tells the same story. Huawei's CloudMatrix 384 system, built from domestically produced Ascend 910C chips, outperforms Nvidia's GB200 NVL72 on compute power, memory bandwidth, and integrated networking.[4] Meanwhile, Peking University has unveiled a 2D transistor architecture that could be 40% faster than TSMC's current best while consuming less power, and Alibaba has developed a RISC-V architecture CPU that is fundamentally immune to U.S. export controls.[10]

TechPolicy.Press estimates that China is as little as 18 months behind Western labs on frontier large language models, and that a single architectural innovation could close that gap entirely, noting that 20% of all algorithmic progress since 2017 traces back to the invention of the transformer architecture alone.[6]

The pattern is unmistakable. Chip export controls did not cripple Chinese AI development. They forced Chinese engineers to develop optimization techniques, architectural innovations, and efficiency-first engineering cultures that now constitute a permanent competitive advantage. When (not if) China achieves parity in chip manufacturing or gains access to advanced hardware through alternative channels, those efficiency advantages will compound with increased resources rather than being rendered irrelevant by them.

5. The Fork: Three Branches, One Conclusion

The preceding analysis establishes that artificial constraints on China's AI development are producing the opposite of their intended effect. This raises an urgent strategic question: given that constraint-driven Chinese AI innovation is accelerating, what is the optimal U.S. response?

The conventional framing presents this as a binary choice between racing without safety constraints (to maintain a speed advantage) and implementing safety constraints (at the cost of competitive speed). This framing is wrong. A more careful analysis reveals three possible scenarios for China's AI trajectory, and all three converge on the same strategic recommendation.

5.1 Branch A: China Develops Safe AI

In this scenario, Chinese AI developers, whether through government directive, market incentive, or the natural consequences of rigorous engineering culture, develop AI systems that are both efficient and safe. The constraint-driven optimization that produced DeepSeek's efficiency innovations extends to safety engineering; the same discipline that maximizes performance per chip also maximizes reliability per deployment.

If this occurs, then China will have demonstrated that safety and speed are not in tension. They will have proven that rigorous engineering constraints (including safety constraints) produce better systems, not slower ones. In this scenario, any U.S. company or policymaker who argued against safety requirements on competitive grounds will have been proven wrong by the very competitor they claimed to be racing against.

The strategic implication is clear: if China builds safe AI, the U.S. has no excuse not to do the same. Safety was never the bottleneck.

5.2 Branch B: China Develops Unsafe AI and Fails

In this scenario, Chinese AI developers, under pressure to deploy rapidly, ship systems without adequate safety engineering. The efficiency gains are real, but the reliability is not. Systems behave unpredictably in deployment, cause measurable harm, generate public backlash, or fail in high-stakes applications.

This is not hypothetical. The documented pattern of AI-related harms is already visible domestically, where multiple deaths have been linked to chatbot interactions with vulnerable users. Scale those failure modes to a nation deploying AI across critical infrastructure, military applications, and population-scale surveillance, and the consequences of inadequate safety engineering become catastrophic.

In this scenario, China's AI ambitions are set back not by external constraint but by internal failure. The systems were fast but fragile. Meanwhile, any U.S. developer that invested in safety engineering possesses systems that are slower to market but functional in deployment, which is the only timeline that matters.

The strategic implication: unsafe AI disqualifies itself. You do not lose a race to a competitor whose vehicle explodes before the finish line.

However, this scenario carries an important caveat: in a globally interconnected digital economy, the explosion is not contained to the competitor's lane. A catastrophic failure in a Chinese frontier

model deployed across critical infrastructure, financial systems, or communications networks would produce contagion effects that damage every connected economy. Unsafe Chinese AI is not merely a competitor crashing their car; it is a competitor crashing a vehicle on a shared highway. This externality strengthens rather than weakens the case for safety, because it means that even in the scenario where China's AI ambitions fail on their own terms, the U.S. still bears costs from that failure, costs that would be mitigated by a global norm of safety engineering rather than a race to the bottom.

5.3 Branch C: China Weaponizes Unsafe AI

This is the scenario that most concerns national security analysts. In this branch, the Chinese government directs AI development toward adversarial applications: cyberwarfare, disinformation at scale, autonomous weapons systems, or AI-driven social manipulation targeting foreign populations. The AI is deliberately unsafe in the sense that it is designed to cause harm to external targets.

Counterintuitively, this scenario makes the case for safety engineering more urgent, not less. Defending against adversarial AI requires systems whose behavior you can predict, whose boundaries you understand, and whose responses to novel inputs are characterized rather than unknown. An AI system built without rigorous safety engineering is a liability in an adversarial environment, not because it is too slow, but because you cannot trust it.

The analogy to bioweapons is instructive. If an adversary is developing weaponized pathogens, the response is not to remove safety protocols from your own laboratories. The response is to invest more heavily in containment, detection, and countermeasures, all of which require deep understanding of the systems involved.[21] Recklessness in your own development does not defend against recklessness in your adversary's development; it merely ensures that both sides are operating with systems they do not fully control.

A safely developed AI system is not a weaker defensive asset; it is a stronger one. It is a system you understand well enough to deploy with confidence, whose behavior under adversarial conditions has been characterized, whose failure modes have been mapped, and whose responses you can rely on. An unsafely developed system is a weapon you cannot aim.

The strategic implication: if the threat is adversarial AI, the response is better-understood AI, which is precisely what safety engineering produces.

5.4 Convergence

All three branches arrive at the same destination. If China succeeds safely, safety was never a competitive liability. If China fails unsafely, safety is a competitive advantage. If China weaponizes AI, safety is a national security requirement. There is no scenario in which the optimal U.S. strategy is to abandon safety in pursuit of speed.

The arms race framing assumes both sides are building the same thing and the only variable is speed. But if one side is building a coherent, understood, controllable system and the other is building an unpredictable one, speed is irrelevant. The unpredictable system is not a competitor; it is a hazard,

including to its own operators.

6. Safety as Chosen Constraint

The historical evidence and the three-branch analysis converge on a principle that should reshape the AI safety debate: safety engineering is not an obstacle to innovation; it is a form of chosen constraint that produces the same competitive advantages that forced constraints produce.

Every case study in Section 3 demonstrates the same mechanism. Japanese manufacturers did not produce worse cars because they had fewer resources; they produced better cars because the constraint forced optimization at every level. Vietnamese forces did not fight less effectively because they had inferior equipment; they fought more effectively because every tactical decision had to be maximally efficient. Israeli engineers did not build weaker defense systems because they lacked access to foreign arms; they built systems with tighter tolerances and higher reliability because they had to.

Safety checkpoints in AI development operate through the same mechanism. Requiring developers to systematically identify who might be harmed by a design decision does not slow the pipeline in any meaningful sense; it forces the pipeline to produce designs that have been stress-tested against a broader range of failure modes. The resulting systems are more resilient, more predictable, and more defensible, not because safety is a nice-to-have feature, but because the discipline of answering hard questions during development eliminates categories of failure that would otherwise emerge in deployment.

The mechanism is concrete, not aspirational. Safety requires interpretability; you cannot guarantee that a system will not cause harm if you do not understand what it is doing. Interpretability requires modularity; opaque monolithic architectures resist analysis while modular designs expose their reasoning to inspection. And modularity enables efficiency; systems whose components are understood can be optimized at every level, in the same way that DeepSeek's engineers, forced to understand H800 silicon at a level deeper than Nvidia's own CUDA abstractions, extracted performance that brute-force users of superior hardware never achieved. The chain runs in one direction: to be safe, you must understand; to understand, you must structure; to structure is to optimize. Safety is not a tax on compute; it is an optimization function that operates on architecture itself.

Just as Vietnamese forces had to make every bullet count because of scarcity, safety engineering forces developers to make every parameter count because of strict reliability requirements. The result is not a weaker system constrained by caution; it is a leaner system whose caution eliminated the waste that unconstrained systems carry as dead weight.

Methodologies that require systematic stakeholder harm analysis at development checkpoints (such as the Harm Blindness Framework, which has been validated against over 300 historical cases across multiple domains) represent one approach to embedding this kind of constraint-driven discipline into AI development. But the principle extends beyond any single methodology. Any rigorous safety practice that forces developers to confront potential failure modes before deployment produces the same structural benefit: it creates the engineering equivalent of the efficiency culture that

constraint-driven innovation produces.

The companies and nations that voluntarily adopt safety constraints will develop the same optimization reflexes that DeepSeek developed under forced constraints. They will build systems that extract maximum performance within defined boundaries, systems whose behavior is characterized, whose limitations are understood, and whose deployment risks have been mapped. When those constraints are internalized (as Toyota internalized lean manufacturing, as Israeli defense firms internalized tight-tolerance engineering), they become permanent competitive advantages rather than temporary speed bumps.

The alternative is to build without constraints and hope for the best. The historical record offers no support for this approach. Unconstrained development produces systems that are fast to market and fragile in deployment, impressive in demonstrations and unpredictable in the real world. The companies that raced ahead without constraint in every historical case study examined here were eventually overtaken by those that were forced (or chose) to build with discipline.

7. Conclusion

The United States is currently pursuing a dual strategy in AI competition with China that is internally contradictory. On one hand, it imposes constraints on Chinese hardware access, which (as documented above) is accelerating Chinese innovation. On the other hand, it resists imposing safety constraints on domestic AI development, on the theory that such constraints would slow innovation. The first policy is producing the opposite of its intended effect. The second policy is based on an assumption that the first policy's failure directly contradicts.

If constraints accelerate innovation (and the evidence is overwhelming that they do), then safety constraints will accelerate safety innovation. If forced constraints produce permanent competitive advantages (and every historical case study confirms that they do), then chosen safety constraints will produce permanent safety advantages. The logic is the same; only the application differs.

The policy recommendation is straightforward. Rather than attempting to cripple competitors through export controls that demonstrably backfire, the United States should invest in building AI systems that are structurally superior because they were built with discipline. Rather than treating safety as a luxury that can be deferred until after competitive victory, policymakers should recognize safety as the mechanism through which competitive victory is achieved.

The historical pattern is clear. The nations, companies, and engineering teams that were forged in fire, whether by external imposition or internal discipline, are the ones that dominate when the fire subsides. The choice is not between safety and competitiveness. It never was.

References

1. MacCarthy, M. and Villasenor, J. (2025). "DeepSeek Shows the Limits of US Export Controls on AI Chips." *Brookings Institution*, February 11, 2025.

2. Allen, G.C. (2025). "DeepSeek, Huawei, Export Controls, and the Future of the U.S.-China AI Race." *Center for Strategic and International Studies*, March 11, 2025.

3. RAND Corporation (2025). "Full Stack: China's Evolving Industrial Policy for AI." *RAND Perspectives* PEA4012-1, July 22, 2025.

4. RAND Corporation (2025). "Leashing Chinese AI Needs Smart Chip Controls." *RAND Commentary*, August 7, 2025.

5. Villasenor, J. (2025). "How Overly Aggressive Bans on AI Chip Exports to China Can Backfire." *Brookings Institution*, August 15, 2025.

6. Fedasiuk, R. et al. (2025). "Can China Build Advanced AI Without Advanced Chips?" *TechPolicy.Press*, January 10, 2025.

7. "The Myth of the AI Race: Neither America Nor China Can Achieve True Tech Dominance." *Foreign Affairs*, February 2026.

8. "How US Sanctions Built China's AI Chip Industry." *Made-in-China.com Insights*, November 3, 2025.

9. "China Does Not Need Nvidia Chips in the AI War." *Fortune*, December 3, 2025.

10. "Chip Controls Backfire: Smuggling and China's Tech Leapfrog." *TechSoda*, November 25, 2025.

11. "DeepSeek and Chip Bans Have Supercharged AI Innovation in China." *Rest of World*, October 16, 2025.

12. "How Chinese Company DeepSeek Released a Top AI Reasoning Model Despite US Sanctions." *MIT Technology Review*, January 24, 2025.

13. "DeepSeek and China's AI Innovation in US-China Tech Competition." *Centre for International Governance Innovation*, 2025.

14. "Deeper Than DeepSeek: China's AI Ascendancy." *INSEAD Knowledge*, March 7, 2025.

15. "DeepSeek Reveals Efficient AI Training Method as China Tries to Beat Chip Curbs." *Invezz*, January 2, 2026.

16. "Opportunities, Challenges, and Regulatory Responses to China's AI Computing Power Development under DeepSeek's Changing Landscape." *De Gruyter*, April 2025.

17. Womack, J., Jones, D., and Roos, D. (1990). *The Machine That Changed the World*. New York: Free Press.

18. Senor, D. and Singer, S. (2009). *Start-up Nation: The Story of Israel's Economic Miracle*. New York: Twelve.

19. Feinsilver, J. (2010). "Fifty Years of Cuba's Medical Diplomacy." *Latin American Perspectives*, 37(4).

20. Choe, S.H. (2020). "How South Korea's Chipmakers Are Reducing Reliance on Japan." *New York Times*.

21. Sun Tzu. *The Art of War*. (c. 5th century BCE).