# Artificial Intelligence and the US-China Balance of Power

by

Benjamin Angel Chang

A.B., School of Public and International Affairs
Princeton University, 2014

Submitted to the Department of Political Science
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Political Science
at the
Massachusetts Institute of Technology

June 2021



Signature of Author ……...
Department of Political Science
June 1, 2021

Certified by ……...
Vipin Narang
Associate Professor of Political Science
Thesis Supervisor

Accepted by ……..
Fotini Christia
Professor of Political Science
Chair, Graduate Program Committee

# Artificial Intelligence and the US-China Balance of Power

by

Benjamin Angel Chang

Submitted to the Department of Political Science at the Massachusetts Institute of Technology on June 1, 2021 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Political Science

## Abstract

How will artificial intelligence affect the US-China balance of power? While a nascent literature debates whether AI may upend strategic stability or revolutionize the nature of warfare, existing discussions suffer from both imprecise conceptualization and scarce data. In three essays, this dissertation evaluates the impact of AI on the nuclear balance, the conventional balance, and long-term US-China competition more generally by focusing on deep learning, generating data through simulation and supply chain analysis.

The first essay defends the focus on deep learning, then presents an end-to-end conceptualization of how its technical qualities translate into usefulness across different categories of modern military tasks, which in turn affect, when contextualized to the particular dyad under study, the strategic balances across different domains of US-China competition. At each analytic layer, the paper condenses deep learning's effects into several generalizations, tying AI to existing debates in security studies and setting an agenda for future research.

The second essay simulates US-China nuclear war in Python to assess AI's impact on the strategic balance, focusing on the tracking of mobile platforms on land. It finds that AI reduces the total "effective counterforce area" – the area the United States would have to destroy with nuclear weapons, to carry out a splendid first-strike – by one to two orders of magnitude. Under low to medium alert, the simulation finds this would enable successful US nuclear counterforce. While countermeasures are available to China, the essay predicts heightened nuclear tensions as a result.

Finally, the third essay exploits supply chain datasets to assess each side's ability to bring AI-enabled autonomous weapons to bear in future conventional conflicts. I find that control over the production of advanced AI chips by the United States and allies almost certainly means the United States would better exploit such weapons, if they emerged as decisive in modern warfare, within at least the next ten years. Potential Chinese policy responses, such as cannibalizing its civilian sector or substituting with older chips, would likely fail for technical reasons.

Thesis Supervisor: Vipin Narang
Title: Associate Professor of Political Science

# Acknowledgments

To borrow from our collective parlance, dissertations never emerge endogenously from graduate students themselves. They are, rather, provoked into existence by confluences, both designed and accidental, of experiences, books, mentors, deadlines, aptitudes, current affairs, daydreams, and, most of all, the intellectual community created by sustained interaction with one's colleagues. I have been exceptionally lucky in this last regard, and to all those without which this dissertation would not be, I am deeply thankful.

First and foremost, I am profoundly grateful to the members of my dissertation committee, Vipin Narang, Chappell Lawson, and Eric Heginbotham. Vipin's sharp instincts and considered guidance on how best to approach dissertating about an emerging technology without stumbling into a quagmire were priceless, while his endlessly good-natured conviviality made the long slogs bearable. This dissertation would not have been attempted in the first place without Chap, who likely changed my life in a single session of straight talk by convincing me to be more ambitious and directly study what most interested me. For this, his always-incisive dissertation comments, and many other favors since, he has my enduring gratitude. Lastly, I owe most of what I know about modeling combat to Eric, who not only educated me about a great many military matters over the years, but also provided a steady stream of invaluable life advice as I worked through the back half of my twenties. I will miss dearly our regular chats in his office.

Outside my committee, the department's world-class intellectual environment made me a better scholar every year. While I learned something in every class and hallway conversation, I am particularly indebted for my education to Barry Posen, Owen Cote, and Taylor Fravel. Collective toleration and critique of my rambling about AI by my peers, for which I will eternally be thankful, sharpened this dissertation over workshops, colloquia, and many a coffee. On these occasions, Erik Sand, Cullen Nutt, and Andy Halterman each provided exceptionally helpful comments. I am further grateful for the deep sense of camaraderie all of us in the department shared throughout the graduate school process, both within and outside my own cohort. I benefited from countless intellectual and social interactions with each of my fellow students, and especially recall stimulating exchanges about philosophy, international affairs, and ethics with Rachel Odell, Sara Plana, and John Minnich, very late nights spent yelling at code with Jesse Clark, Gabriel Nahmias, and Emilia Simison, and a constellation of other rich experiences besides. Outside class, Joli Divon Saraf and Susan Twarog provided vital logistical support which made everything outside learning itself easy and non-distracting. Attending MIT was an immeasurable privilege.

Beyond MIT, this dissertation also owes a great deal to my colleagues at Georgetown University's Center for Security and Emerging Technology, where perhaps the most intense study of AI's policy implications is being undertaken today. To the degree that my knowledge about AI contains original thought, it is a reflection of the intellectual environment created by Jason Matheny and now ably stewarded by Dewey Murdick, Lynne Weil, Tessa Baker, Helen Toner, Igor Mikolic-Torreira, and many others. I have benefited inexpressibly from mentorship on all matters by Matt Daniels, while Ben Buchanan, Saif Khan, and Will Hunt each provided critical advice on the research direction of the dissertation. Andrew Imbrie, Tarun Chhabra, and Melissa Flagg have proved inexhaustible sources of guidance on both AI and non-AI matters alike, and I further owe large swaths of my thinking about China and AI to conversations with Remco Zwetsloot, Dahlia Peterson, Anna Puglisi, Carrick Flynn, Ben Murphy, Jeff Alstott, Tim Hwang, Roxanne Heston, Tina Huang, Rebecca Kagan, and many others both within and outside CSET too numerous to name. For all those who have generously served as intellectual interlocutors of mine in DC, I can only hope that they will forgive my poor translation of their ideas into academic-ese.

From other institutions, I have received generous support from the National Science Foundation through its Graduate Research Fellowship Program, which provided funding for my graduate studies under Grant No. 1122374. Of course, any opinions expressed herein do not necessarily reflect the views of any other people

or organizations, the National Science Foundation or otherwise. An astonishingly efficient education in nuclear weapons modeling by Daryl Press, Keir Lieber, Brendan Green, and others at the Strategic Force Analysis Boot Camp at Sandia National Laboratories also informs this dissertation's second chapter.

Finally, without Torin Rudeen, a lifelong friend, actual AI expert (as opposed to us pale imitators outside computer science), and my co-author on the second paper below, this dissertation simply would not have happened. He served as critical sources of both technical knowledge and moral support through not only the writing of the entire dissertation, but also throughout the whole PhD, and I am profoundly grateful.

I must conclude on a somber note. During my time in graduate school, both my father and my grandmother passed away. They had high hopes that I would use my many privileges to leave the world a better place than I found it, and I intend to do my best to satisfy those hopes. I dedicate this dissertation to their memory.

# Table of Contents

# Introduction

What is artificial intelligence, and how will it affect our world? Perusing the existing literature, one gradually acquires the impression that AI may revolutionize warfare, obsolete all human labor, cast the globe into permanent authoritarianism, and/or more generally either destroy our world or save it. Stephen Hawking, for example, famously warned that "full artificial intelligence could spell the end of the human race."[1] Similarly, for Henry Kissinger, AI could mean the end of the Enlightenment, causing human history to "go the way of the Incas."[2] Even short of the end of the human age, for Graham Allison, China could use AI to overtake US military power by obsolescing large swaths of our advanced but still human-reliant forces, achieving a kind of "AI supremacy."[3] Alternatively, as one reads further about AI, it may become difficult to avoid the slight suspicion, kept in the back of the mind, that perhaps AI is "just BS," a passing fad likely to leave our theoretical understanding of world affairs largely unscathed.[4]

What are we to make of all this? In my view, the present conceptual morass arises from this issue: there is no consensus definition of AI. Consequently, those writing within security studies (and in adjacent, non-academic venues) are commenting on a large, loosely constrained grab-bag of different technologies, some mature, some newly born into the world, and some not yet existent. Undisciplined by technical bounds or the empirics generated by mature military technologies with a battlefield record, amateur speculation about AI has essentially no bounds: writers are free to use AI as a blank slate onto which to project their particular worries, whether that be the amplification of structural inequities in American society, the military and ideological global struggle against Chinese Communism, the intensifying disempowerment of blue-collar workers due to industrial automation, or even the wholesale destruction of the human race.

To be clear, I do not mean to suggest that all or even any of these narratives are wrong – in fact, each of these worries has generated a particular, genuine scholarly community based outside political science, discussing the impact of AI on each concern in earnest. Rather, I mean that among observers of those affairs related to international security (whether in universities or at think tanks), extant discussions have been, for the most part, conceptually undisciplined, drawing on and combining existing strands of thought from other conversations about, in fact, very different technologies and applications. This has resulted in confusion.

## Defining Artificial Intelligence

Thus, this dissertation argues that scholars in security studies should mainly study *deep learning*, a specific kind of machine learning leveraging multiply-layered "neural networks," which are computing constructs loosely inspired by the human brain. Among technologies commonly called AI, deep learning can be contrasted with symbolic AI, a first-wave technical approach most prominent beginning in the 1950s, and artificial general intelligence (AGI), a hypothetical future capability with AI as good as humans at all tasks.

In my view, focusing on deep learning captures, among technologies thought of as AI, the effects most important to security studies, such as the possibility of future "drone swarms" of perhaps thousands of

[1] Rory Cellan-Jones, "Stephen Hawking warns artificial intelligence could end mankind," BBC, 2014, https://www.bbc.com/news/technology-30290540.

[2] Henry A. Kissinger, "How the Enlightenment Ends," *The Atlantic*, June 2018, https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

[3] Graham Allison, "Is China Beating America to AI Supremacy?," *The National Interest*, December 22, 2019, https://nationalinterest.org/feature/china-beating-america-ai-supremacy-106861.

[4] Daniel W. Drezner, "What if AI is just BS?," *The Washington Post*, May 1, 2019, https://www.washingtonpost.com/outlook/2019/05/01/what-if-ai-is-just-bs/?

autonomous weapons platforms, discussed in this dissertation's third essay, or the potential for AI-assisted nuclear counterforce, discussed in this dissertation's second essay. Many papers could – and have – been written about symbolic AI and AGI, but they cannot be productively conflated with deep learning.

| Symbolic AI | Deep Learning | Artificial General Intelligence |
|---|---|---|
| • **Technical Basis:** physical symbol systems<br>• **Expert Systems**<br>(e.g., tax prep software, medical diagnostic aids)<br>• **Feedback Control Systems**<br>(e.g., aircraft autopilots, Perimeter) | • **Technical Basis:** multiply-layered neural networks<br>• **Pattern Recognition**<br>(e.g., facial recognition, the protein-folding problem)<br>• **Pattern Generation**<br>(e.g., deepfakes, human-indistinguishable writing) | • **Technical Basis:** unknown – various candidates<br>(e.g., whole brain emulation, future self-modifying machine learning, brain-computer interfaces, etc.)<br>• **All Human Functions**<br>(e.g., dissertation-writing) |

What narrowing the focus to deep learning reveals is that despite the broad generalities characterizing much existing commentary, AI's effects on variables of interest to security studies will likely vary by both domain and dyad. First, variation by domain emerges because deep learning's contribution to different military tasks will likely be fairly uneven. For missions like electronic warfare built atop mature engineering disciplines, for example, new AI algorithms must clear a higher bar to improve upon the state-of-the-art. In contrast, since previous-generation, manually-crafted systems struggled with characterizing images and understanding language, deep learning's applications to computer vision and automated intelligence-processing represent relatively large gains.[5] AI is not itself a discrete physical object, like machine guns, fighter aircraft, or nuclear weapons, but rather a general-purpose technology with many applications; consequently, its manifestations importantly vary across domains in their effects.

Second, variation by dyad emerges because AI's strategic effects are, I argue throughout this dissertation, mediated both by each country's stock of AI's *inputs* – those raw precursors, including data, hardware, and engineering talent, necessary to create and deploy AI – as well as each country's *leverage*, or their ability to use AI applications, once produced, to achieve strategic goals specific to the country. Since technologies are often more useful for some goals than others, differences in leverage can be decisive. This dissertation thus focuses on the US-China balance, both due to the dyad's substantive importance, as well as because both countries are among the farthest along in deploying AI.

## Artificial Intelligence and the US-China Balance of Power

Overall, contrary to the conventional wisdom that China may use AI to leapfrog US national power, I find that AI will more likely strengthen the US position relative to China in both the nuclear and conventional domains. More generally, as opposed to the received wisdom about information technology in the age of globalization, I find that AI's effects likely drive "rich get richer effects," as opposed to democratizing access to military power. The United States is advantaged against China in leveraging AI, and in turn, China is advantaged against its own domestic population. In both cases, the more well-resourced agent wins.

In the nuclear domain, AI may counterfactually enable a successful US nuclear counterforce campaign against the Chinese arsenal. Leveraging a detailed model in Python coded with Torin Rudeen, the paper reports the results of simulating US counterforce both with and without AI, and finds that AI reduces the total "effective counterforce area" – the area the United States would have to destroy with nuclear weapons, to carry out a splendid first-strike – by one to two orders of magnitude. Across a variety of stochastic and selected assumptions, this makes the difference in whether a US first-strike is successful.

[5] Tom Stefanick, "AI in the Aether: Military Information Conflict," in *The Global Race for Technological Superiority*, ed. Fabio Rugge (Brookings, 2019), https://www.brookings.edu/wp-content/uploads/2019/12/FP_20191211_military_information_conflict_stefanick.pdf, 112-30.

In the conventional domain, while it remains an open question whether fully autonomous AI-enabled drones will revolutionize future warfare, collected in swarms or otherwise, I exploit industry datasets to argue that control over the production of advanced AI chips by the United States and allies almost certainly means that such drones would not, if they emerge, advantage China. Based on longstanding technical trends, I argue that AI weapons will require cutting-edge chips – China cannot substitute with large numbers of older chips. I evaluate various possible Chinese policy responses, up to and including an invasion of Taiwan to seize its chip production capabilities, and find that they would likely fail to restore Chinese access in the event of a US-led supply cutoff. Thus, if the future of warfare comes to be dominated by "drone swarms," this will favor the United States at least for the next ten years, and likely afterwards as well.

## Artificial Intelligence and Security Studies

This dissertation has a number of implications for scholars of security studies. First, for those interested in artificial intelligence, the first essay offers the field an operationalization of AI as deep learning, providing an object amenable to study by accepted methods in security studies. The second and third essays, which deploy campaign analysis and supply chain analysis, respectively, provide methodological examples of how to study an emerging technology not yet deployed in fully realized form on the world's battlefields. The third essay also advances the discussion of drone swarms by providing technically grounded theory about how they would work, and what they would materially require.

Second, for nuclear scholars, the second essay represents quasi-experimental evidence about the effects of an emerging technology, AI, on the general survivability of nuclear arsenals. Mobile delivery platforms, both on land and at sea, exploit the simple fact that, to put it colloquially, nations and oceans are large, while launchers and submarines are small; consequently, mobility has represented a strong form of concealment owing to difficulties in locating moving platforms. Our model, however, shows that with AI, mobility may no longer be a guarantee of obscurity. States have several options for countermeasures, which the second essay explores in detail. Many of these countermeasures, however, such as seeking to degrade adversary AI systems, destroying adversary space assets, or switching to a launch-on-warning posture, may increase the chances of escalation, both accidental and intended. While our model focuses on the US-China balance, Iran and North Korea are likely to be even worse off vis-à-vis the United States, and states in other dyads – such as India, considering its nuclear strategy against Pakistan – may feel tempted to explore AI-enabled counterforce options as well.

Third, for students of warfare, the third essay gives technical reasons grounded in the political economy of transnational production to show that the emerging AI capabilities in conventional warfare most often speculated about will likely lean on a critical material precursor – advanced AI chips – whose production is controlled by the United States and its allies. "AI chips," I argue, are analogous to highly enriched uranium in the early nuclear age – while uranium occurs plentifully and naturally in nature, and chips are freely sold in iPhones, laptops, and even cars, the ability to produce the specialized chips required for AI weapons will serve as a key supply-side constraint – and source of US advantage – for at least ten years.

Fourth, for those interested in the diffusion of military power, the industry datasets exploited in the third essay show that globalization, rather than leading to redundancy and the reduced ability of any given state to impose supply cutoffs on others, has in fact led to specialization and the consequent development of monopolies, or near-monopolies, at various high-tech chokepoints. Put another way, contrary to cyclic theories beginning with Robert Gilpin which argue that hegemons invariably undermine themselves by encouraging technological diffusion to the periphery, I find that in production of cutting-edge technologies in particular, R&D intensity may prevent diffusion and instead re-concentrate power in the leading state.[6]

[6] Robert Gilpin, *War and Change in World Politics* (Cambridge: Cambridge University Press, 1981), 162, 176-85.

## Artificial Intelligence and US Policymaking

This work also has implications for policymakers. First, some observers have argued that because China may overtake the United States by means of AI, we should dispense with excessive concerns about ethics and AI misbehavior and instead focus on catching up. This is backwards. AI is instead likely to augment US power vis-à-vis China, if anything, and making sure AI technologies are safe for the battlefield both honors longstanding US values and builds trust necessary to effective organizational adoption of AI by the US military – if soldiers do not trust the systems into whose artificial hands they must put their lives, this will hurt, rather than enhance, US combat effectiveness.[7]

Second, policymakers must maintain strategic stability with China to avoid the risk of accidental nuclear war. While China has thus far resisted nuclear arms control talks with the United States, avoiding trilateral discussions with Russia, AI counterforce capabilities may eventually compel greater diplomatic engagement. US policymakers should not stumble into an AI-empowered counterforce capability, but rather undertake a strategic review and deliberately choose, with eyes wide open, whether or not to hold the Chinese arsenal at risk. Simultaneously, policymakers must attend to related risks in other dyads, such as with India and Pakistan, where AI could similarly tilt the scale of enabling (or appearing, to policymakers, to enable) counterforce.[8]

Finally, this work offers several conceptualizations, reached in the course of its three essays, of how best to compete with China in AI. Being ahead in AI *research* lacks inherent virtue – famously, the French first invented mechanically powered submarines, but regretted this when the British Navy, better adopting the technology, leapfrogged French sea power.[9] While AI research will invariably diffuse, however, AI *hardware* chokepoints are a strong, weapons-relevant advantage worth maintaining. Similarly, as the first essay details, the United States has a longstanding advantage over China in attracting high-skilled, AI-relevant STEM talent it would be foolish to discard. The United States should specifically avoid advancing research which, once diffused, would asymmetrically empower China to reach its goals – for example, AI applications enabling information warfare most effective against democratic states which hold elections. Instead, the United States should coordinate with its allies to establish robust AI technology standards which favor democracies globally.

## Dissertation Outline

The above-synthesized analysis unfolds in three essays below. Summarizing, the first essay asks, "What is artificial intelligence, and what are its effects?" It argues that the extant literature suffers from a badly operationalized independent variable – as scholars define AI in very different ways, analyses disagree simply due to discussing entirely different technologies. This brushclearing paper operationalizes AI as deep learning, then presents an "end-to-end" conceptualization of deep learning's causal impact, moving from its basic technical nature ("mechanical effects") to its usefulness across different categories of modern military tasks ("tactical effects"), then finally to its contextual impact on the US-China balance of power ("strategic effects"). At each analytic layer, the paper condenses deep learning's effects into several generalizations, tying AI to existing debates in security studies and setting an agenda for future research.

---

[7] Michèle A. Flournoy, Avril Haines, and Gabrielle Chefitz, "Building Trust through Testing: Adapting DOD's Test & Evaluation, Validation & Verification (TEVV) Enterprise for Machine Learning Systems, including Deep Learning Systems," *WestExec Advisors*, 2020, https://cset.georgetown.edu/wp-content/uploads/Building-Trust-Through-Testing.pdf.

[8] Christopher Clary and Vipin Narang, "India's Counterforce Temptations: Strategic Dilemmas, Doctrine, and Capabilities," *International Security* 43.3 (2019), 7-52.

[9] Michael C. Horowitz, *The Diffusion of Military Power: Causes and Consequences for International Politics* (New Jersey: Princeton University Press, 2010), 1-2.

The second paper, coauthored with Torin Rudeen, asks, "Will artificial intelligence threaten the survivability of nuclear arsenals?" To assess this question, we code a simulation of US nuclear counterforce efforts against the Chinese arsenal, separately modeling each of "find, fix, and finish." We run this simulation both with and without AI at various alert levels, finding that AI threatens the survivability of Chinese road-mobile missiles at low and medium alert. Our simulation represents the first detailed modeling of the question of mobile platform detectability from first principles, as well as a significant advance in model sophistication over the open-source state-of-the-art. On balance, our results provide evidence for the view that technological developments may increasingly threaten strategic stability.

Finally, the third essay seeks to make inroads on the question, "Will artificial intelligence favor China or the United States in the conventional military domain?" It argues that specialized AI chips will be the key material bottleneck for states' ability to field AI-enabled weapons platforms at scale. Critically, while AI research is open, global, and collaborative, AI *hardware* is ever more consolidated in US and allied countries. The involved technologies are approaching fundamental physical limits; consequently, leading firms require both almost all available global talent, as well as the economy of scale created by meeting almost all global demand, to make continuous R&D progress. Thus, for deep structural reasons related to long-term trends in science, I argue that AI chip production capability will not diffuse to China for at least the next ten years, giving the United States a fundamental material advantage in AI-enabled conventional weapons for that period and likely beyond.

Conceptually, the collected logic of the dissertation can be understood as follows: at a high level, the first paper identifies *search* and *mass* as main effects of AI on military power. The second paper applies *search* to the *nuclear* domain, using simulation to assess AI's impact on the US-China strategic balance through AI-empowered tracking of mobile platforms. Complementarily, the third paper applies *mass* to the *conventional* domain, exploiting proprietary datasets related to the AI supply chain to assess each side's ability to bring mass in the form of AI-enabled weapons to bear in future conventional conflicts.

# On Not Confusing Ourselves About Artificial Intelligence: Deep Learning, Security Studies, and the US-China Balance of Power

## Introduction

What is artificial intelligence, and what are its effects? Recent breakthroughs in modern AI have sparked fierce attention from the world's national security ecosystems: Putin has averred that "whoever leads in AI will rule the world," and most countries of geopolitical note have developed national AI policies, appearing to share his urgency.[10] China's "Next Generation Artificial Intelligence Development Plan" assesses that AI, as a "new focus of international competition," is a "technology that will lead in the future"; the plan calls for China to "lead the world in AI" by 2030.[11] According to the US National Security Commission on Artificial Intelligence, AI weapons will be "weapons of first resort in future conflicts" and "transform all aspects of military affairs." On their view, Chinese strategy seeks specifically to "leapfrog" US power by more quickly adopting AI, which it sees as underpinning the next generation of military technology.[12]

This flurry of activity has helped provoke, in turn, an equally sizable crush of commentary. Prominent takes disagree wildly: is AI "the new electricity," to use the phrase Andrew Ng famously coined, or very possibly "just BS" – today's nanotechnology – as Daniel Drezner muses?[13] Observers variously speculate that AI may automate perhaps 40% of human jobs in the near-term, revolutionize the nature of warfare, or even cause human extinction.[14] For Henry Kissinger, AI will bring about the end of the world created by the

[10] "'Whoever leads in AI will rule the world': Putin to Russian children on Knowledge Day," *Russia Today*, September 1, 2017, https://www.rt.com/news/401731-ai-rule-world-putin. For an overview of national AI plans, see Jessica Cussins Newman, "Toward AI Security: Global Aspirations for a More Resilient Future," *Center for Long-Term Cybersecurity*, February 2019, https://cltc.berkeley.edu/wp-content/uploads/2019/02/Toward_AI_Security.pdf, as well as Tim Dutton, Brent Barron, and Gaga Boskovic, "Building an AI World: Report on National and Regional AI Strategies," *CIFAR*, 2018, https://www.cifar.ca/docs/default-source/ai-society/buildinganaiworld_eng.pdf.

[11] China State Council, "A Next Generation Artificial Intelligence Development Plan," July 20, 2017, translated by New America, https://www.newamerica.org/documents/1959/translation-fulltext-8.1.17.pdf

[12] Eric Schmidt, Robert Work, Safra Catz, Eric Horvitz, Steve Chien, Andrew Jassy, Mignon Clyburn, Gilman Louie, Chris Darby, William Mark, Kenneth Ford, Jason Matheny, Jose-Marie Griffiths, Katharina McFarland, and Andrew Moore, "Final Report," *National Security Commission on Artificial Intelligence*, March 2021, https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf, 1, 22-3.

[13] Andrew Ng, "Artificial Intelligence is the New Electricity," *Stanford Graduate School of Business*, 2017, https://www.youtube.com/watch?v=21EiKfQYZXc; Daniel W. Drezner, "What if AI is just BS?," *The Washington Post*, May 1, 2019, https://www.washingtonpost.com/ outlook/2019/05/01/what-if-ai-is-just-bs/?

[14] The "nature of warfare" remark is attributable to former Deputy Secretary of Defense Robert Work. See Benjamin Jensen, Scott Cuomo, and Chris Whyte, "Wargaming with Athena: How to Make Militaries Smarter, Faster, and More Efficient with Artificial Intelligence," *War on the Rocks*, June 5, 2018, https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries smarter-faster-and-more-efficient-with-artificial-intelligence/. The automation remark is from AI entrepreneur Kai-Fu Lee. See Don Reisinger, "A.I. Expert Says Automation Could Replace 40% of Jobs in 15 Years," *Fortune*, http://fortune.com/2019/01/10/automation-replace-jobs/. Opinions are mixed on automation. See Erin Winick, "Every study we could find on what automation will do to jobs, in one chart," *MIT Technology Review*, January 25, 2018, https://www.technologyreview.com/s/610005/every-study-we-could-find-on what-automation-will-do-to-jobs-in-one-chart/. On extinction, see a skeptical view at Oren Etzioni, "No, the Experts Don't Think Superintelligent AI is a Threat to Humanity," *MIT Technology Review*, September 20, 2016, https://www.technologyreview.com/s/602410/no-the-experts-dont-think-superintelligent-ai-is-a-threat-to-humanity/, then a rebuttal at Allan Dafoe and Stuart Russell, "Yes, We Are Worried About the Existential Risk of Artificial Intelligence," *MIT Technology Review*, November 2, 2016, https://www.technologyreview.com/s/602776/yes-we-areworried-about-the-existential-risk-of-artificial-intelligence/. For the seminal work on risks from very advanced AI,

Enlightenment, as human cognition will become increasingly unable to keep up with, and resist, the direction induced in society by AI.[15] In his words, "Humanity is at the edge of a revolution driven by artificial intelligence."[16] For Jason Lanier and Glen Weyl, two researchers at Microsoft, AI "is best understood as a political and social ideology," one which seeks to "replace, rather than complement, not just individual humans but much of humanity." In their view, the idea of AI should be abandoned entirely, in favor of renewed focus on human beings.[17]

Reading the literature, in short, one gets the mounting impression that AI may lead to the final victory of Chinese Communism, make nuclear arsenals everywhere vulnerable to enormous swarms of autonomous drones, destroy humanity altogether, and/or simply be the latest fad in buzzwords among those whose profession depends on the continuous discussion of such buzzwords, in the same mental category as "Big Data" or "globalization" or "5G."[18] Insofar as something important may be happening here, the mind nonetheless begins to desire greater precision: again, "what is AI, and what are its effects?"

We are, I believe, confusing ourselves. Generating useful analysis of new technologies requires especially precise conceptualization, since the empirics which would typically discipline errant speculation are scarce. Since common understandings of what AI is range both tax software and weaponized drones, asking "what are the effects of AI?" without proper operationalization is like asking, "what are the effects of explosives?", lumping together firecrackers and nuclear weapons. The analytic consequences of persistent conceptual fuzziness are real – for example, one recent article suggesting America adopt an AI-enabled "Dead Hand" conflates three highly distinct technologies.[19] If AI is anything close to as big a deal as some like Kissinger believe, we must do better.

This paper attempts to do better. First, I grapple with the problem of defining AI. I argue that "AI" is best decomposed into several different technologies, each superficially similar but best conceptualized as importantly separate by political science, similar to how interstate wars and civil wars are understood to have distinct causes, characteristics, and effects. I further argue that security studies should, at present, mostly study "deep learning," that approach to machine learning which uses multiple layers of artificial neural networks, which are computing constructs loosely inspired by the brain, to learn how to best represent data. Deep learning underlies most recent headline-grabbing progress in AI.

---

see Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford University Press: UK, 2014). The best quick overview of these concerns is Kelsey Piper, "The case for taking AI seriously as a threat to humanity," *Vox*, December 23, 2018, https://www.vox.com/future-perfect/2018/12/21/18126576/ai-artificial-intelligence-machine-learning-safety-alignment.

[15] Henry A. Kissinger, "How the Enlightenment Ends," *The Atlantic*, June 2018, https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

[16] Henry A. Kissinger, Eric Schmidt, and Daniel Huttenlocher, "The Metamorphosis," *The Atlantic*, August 2019, https://www.theatlantic.com/magazine/archive/2019/08/henry-kissinger-the-metamorphosis-ai/592771/.

[17] Jaron Lanier and Glen Weyl, "AI is an Ideology, Not a Technology," *WIRED*, March 15, 2020, https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/.

[18] Some of my colleagues would virulently disagree with me about 5G, and indeed, my point is not that these terms mean and refer to nothing, but simply that our discussions about them are overwhelmingly imprecise.

[19] Adam Lowther and Curtis McGiffin, "America Needs a 'Dead Hand,'" *War on the Rocks*, August 16, 2019, https://warontherocks.com/2019/08/america-needs-a-dead-hand/. The Soviet Union's Perimetr system relied on the if-then symbolic logic most often called AI beginning in the 1950s, not modern machine learning. Artificial General Intelligence (AGI) is a theoretical possibility about which experts disagree – it may be centuries away, or impossible. Neither is deep learning, the AI approach most responsible for the recent upswell of interest.

Second, I ask what AI's effects are. I argue that operationalizing AI as deep learning clarifies presently confused debates about AI's effects – while no satisfyingly final answers can be reached within one paper, delving into deep learning's particulars helps make AI "legible" to the theoretical constructs of political science, establishing a basis for further research. This second effort occurs in three parts, split into an exploration of deep learning's *mechanical*, *tactical,* and *strategic* effects:

> ● By "mechanical effects," I mean what deep learning does, technically speaking, in a vacuum, or when run on a laptop in the woods, so to speak. Analogously, oil's mechanical effects derive from its nature as a dense hydrocarbon.
>
> ● By "tactical effects," I mean what deep learning does, in the context of its effects on a modern military. Analogously, oil's mechanical effects translated into much greater mobility and duration of maneuver for militaries than before.
>
> ● Finally, by "strategic effects," I mean what deep learning does, in the context of geopolitical competition between the United States and China – that is, what is its effect on the balance of power? Analogously, oil drastically enhanced the power of Saudi Arabia, permanently altering the geopolitics of the Middle East.

To preview in summary form, at the 10,000-foot level we can understand deep learning as, *mechanically*, a technology which enables machines to learn, and consequently both recognize and reproduce, patterns from data. Here, I mean "patterns" in a very powerfully general sense, ranging from "what it looks like when stocks are about to go up," to "what kinds of imagery data human beings would predictably call a cat." *Tactically*, deep learning's pattern-learning abilities permit, for modern militaries, a kind of fungibility of resources: successful adoption allows converting the triple bundle of compute, task-specific data, and AI talent into the automation of a wide variety of military tasks, tasks which either previously required human beings, or were outright impossible due to human limitations. Bucketing, we can usefully understand these tasks as falling significantly into two conceptual categories: *search*, meaning the ability to find proverbial needles in haystacks of various kinds, and *mass*, meaning the converse ability to produce and bring to bear haystacks of useful needles. At both the mechanical and tactical layers, however, technical risks specific to deep learning generate the possibility of new types of accidents and vulnerabilities, complicating adoption.

*Strategically*, translating upwards to how this impacts the US-China balance of power, technologies definitionally alter relative balances of power when they disproportionately benefit, or disproportionately harm, one state versus another. When I embed the above-theorized tactical effects into the US-China context and ask what results, I find that there are differences not only in how much compute, data, and talent either country can acquire (thus possessing the key *resources* which are "*inputs"* to AI), but also in how well each country can use the resulting automation of search and mass (thus *leveraging* the *"outputs"* of AI) across various domains of competition, both nuclear and conventional, to gain relative advantage. As a first cut, I find that AI may worsen the nuclear balance for China, generate mixed effects on the conventional balance, and strengthen China's prospects in long-term peacetime competition with the United States. As no single paper can satisfactorily answer questions across every domain, I primarily aim to lay down a conceptual architecture for further research into the strategic layer of AI's effects.[20] Finally, I conclude the paper by briefly re-summarizing my argument and suggesting avenues and methods for future research.

[20] The second paper of my dissertation further explores the nuclear balance, and the third paper of my dissertation further explores the conventional balance.

## What is AI?

AI has no consensus definition.[21] Varying definitions across the US government cover, rather chaotically, almost all computing-related activities. Several bodies, including the Office of the Director of National Intelligence (ODNI) and the NSCAI, roughly define AI as using computers to replace human cognition.[22] Elsewhere, definitions are far more general – for example, the National Institute of Standards and Technology (NIST) indicates AI is the "capability of a system to acquire, process, and apply knowledge," and Section 238 of the 2019 National Defense Authorization Act (NDAA) gives a maximally inclusive, five-part definition of AI, even including artificial systems developed in neither software nor hardware, including "set[s] of techniques."[23]

Thankfully, matters are somewhat tamer among computer scientists, who mostly form a 2x2 in disagreeing whether intelligence means "human-like" or "ideally rational," plus whether the AI's "thought" or "behavior" must exhibit intelligence.

**Most AI Definitions Within Computer Science Fall Into a 2x2**[24]

| | | **Dimension of Behavior** | |
|---|---|---|---|
| | | Thinking | Acting |
| **Key Criterion for "AI-Ness"** | Human-Like | "machines that think like humans" | "machines that act like humans" |
| | Ideally Rational | "machines that think rationally" | "machines that act rationally" |

[21] This issue is oft discussed. See Daniel S. Hoadley and Kelley M. Sayler, "Artificial Intelligence and National Security," *Congressional Research Service*, January 30, 2019, https://fas.org/sgp/crs/natsec/R45178.pdf, 1-4; Danielle C. Tarraf, William Shelton, Edward Parker, Brien Alkire, Diana Gehlhaus Carew, Justin Grana, Alexis Levehahl, Jasmin Leveille, Jared Mondschein, James Ryseff, Ali Wyne, Dan Elinoff, Edward Geist, Benjamin N. Harris, Eric Hui, Cedric Kenney, Aydne Newberry, Chandler Sachs, Peter Schirmer, Danielle Schlang, Vicotria Smith, Abbie Tingstad, Padmaja Vedula, and Kristin Warren, "The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations," *RAND*, 2019, https://www.rand.org/pubs/research_reports/RR4229.html, 147-154; P. M. Krafft, Meg Young, Michael Katell, Karen Huang, and Ghislain Bugingo, "Defining AI in Policy versus Practice," *arXiv*, December 23, 2019, https://arxiv.org/pdf/1912.11095v1.pdf; Jonas Schuett, "A Legal Definition of AI," *arXiv*, September 4, 2019, https://arxiv.org/pdf/1909.01095.pdf; and Heather M. Roff, "The frame problem: The AI 'arms race' isn't one," *Bulletin of the Atomic Scientists* 75.3 (2019): 95-8, among others.

[22] "The AIM Initiative: A Strategy for Augmenting Intelligence Using Machines," *Office of the Director of National Intelligence*, January 16, 2019, https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf, 13; Eric Schmidt, Robert O. Work, Safra Catz, Steve Chie, Mignon Clyburn, Christopher Darby, Kenneth Ford, Jose-Marie Griffiths, Eric Horvitz, Andrew Jassy, Gilman Louie, William Mark, Jason Matheny, Katharina McFarland, and Andrew Moore, "Interim Report," *National Security Commission on Artificial Intelligence*, November 2019, 7. Available online: https://www.epic.org/foia/epic-v-ai-commission/AI-Commission-Interim-Report-Nov-2019.pdf. This was also the approach of DoD's 2018 AI Strategy, which defined AI as "the ability of machines to perform tasks that normally require human intelligence." (This definition had no legal force. There is no DoD-wide definition.) See "Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity," *DoD*, 5. Available online: https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

[23] Tarraf et al., "The Department of Defense Posture for Artificial Intelligence," 150-2; *National Defense Authorization Act for Fiscal Year 2019*, Section 238, https://www.congress.gov/115/bills/hr5515/BILLS-115hr5515enr.pdf.

[24] Taxonomy from Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (US: Prentice Hall, 2009), 5. This is the standard introductory textbook for computer scientists studying AI.

Confronting these options, security scholars have mostly ticked the "machines that act like humans" quadrant.[25] Within that 2x2, this is the logical choice – after all, behavior is the causally important attribute for weapons systems; it matters not if the AI "thinks." Similarly, in considering when, whether, and how AI will replace human operators at various tasks, "human-like" qualities are the key threshold – since human behaviors are themselves riddled with various cognitive biases, AI need not be "ideally rational" to meet the bar.[26]

In my view, however, no quadrant of this 2x2 best serves our purposes. Critically, the definitions therein are all *outcome-based* – that is, they define as AI all machines which have the ability to exhibit a certain form of *performance*, such as human-like behavior or thinking. Such definitions make a great deal of sense for computer science, which has the creation and refinement of AI as its goal, but not for political science.

First, outcome-based definitions inherently under-exploit the technical literature. Since these operationalizations define something as AI *if* capable of human-like behavior, and not, otherwise, they black-box whether AI can actually produce such behaviors. It is definitionally assumed. Besides risking separation from technical realities, this approach thus also tends to elide sober discussion of AI's limitations and vulnerabilities. Regrettably, much strategic analysis is still patterned, "well, if AI can do X, then Y might follow," without answering, "can AI do X?" In such cases, replacing "AI" with "magic" would not cause loss of content.

Second, more worryingly, outcome-based definitions lack intercoder reliability. Since analysts tend to have different opinions about what "meets the bar" of human-like performance, discussions can give the impression that both everything and nothing can be AI. Arguably, after all, calculators, drone swarms, and tax preparation software all exhibit "human-like behavior" in one sense or another, as they replace counterfactual expenditures of human labor and cognition. As M. L. Cummings notes, "by this definition a house thermostat is intelligent because it can perceive and adjust the temperature."[27] On the other hand, many tasks previously considered benchmarks for "real" AI, such as playing chess or writing poetry, are often retroactively dismissed as "not really requiring *intelligence*" once computers accomplish such tasks. In Douglas Hofstadter's now standard misquoting of the famous adage, "AI is whatever hasn't been done yet."[28] This conceptual confusion is not only deeply unsatisfying, but also precludes analytic progress in studying AI as an independent variable.

---

[25] M. L. Cummings, "Artificial Intelligence and the Future of Warfare," *Chatham House*, January 2017, https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf, 4; Vincent Boulanin, "Introduction," in *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume I: Euro-Atlantic Perspectives*, ed. Vincent Boulanin (SIPRI, May 2019), https://fas.org/sgp/crs/natsec/R45178.pdf; Michael Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1 (2018), https://doi.org/10.15781/T2639KP49; Michael C. ----, "When speed kills: Lethal autonomous weapon systems, deterrence and stability," *Journal of Strategic Studies* 42 (2019), 767.

[26] Molly Kovite, "I, Black Box: Explainable Artificial Intelligence and the Limits of Human Deliberative Processes," *War on the Rocks*, July 5, 2019, https://warontherocks.com/2019/07/i-black-box-explainable-artificial-intelligence-and-the-limits-of-human-deliberative-processes/. Listing human foibles is the standard counter to accident-based worries about deploying AI, imperfections and all. For the seminal work on human biases, see Daniel Kahneman, *Thinking, Fast and Slow* (US: Farrar, Straus and Giroux, 2011).

[27] Cummings, "Artificial Intelligence and the Future of Warfare," 6.

[28] Douglas R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid* (US, Basic Books, 1979), 601. Tesler maintains the quote was "Intelligence is whatever machines haven't done yet," though the Hofstadter rendering is now omnipresent. Larry Tesler, "CV: Adages & Coinages," http://www.nomodes.com/Larry_Tesler_Consulting/Adages_and_Coinages.html.

For example, political science may wish to ask, "does AI increase or decrease X?" for various X (e.g., the cost of air power, the sustainability of authoritarianism, the survivability of nuclear arsenals, and so on). Questions of this sort are ill-formed, however, if no consensus exists on what actually counts as AI, and especially if different subsets of "AI" produce oppositely signed effects. To illustrate, one prominent discussion concerns whether and when AI-driven automation will obsolete different kinds of human experts. Here, however, while expert systems (first created in the 1970s, but still used today) lean on human specialists to formulate enormous databases of manually inputted knowledge, deep learning models often leverage no human domain knowledge whatsoever. The first expert system, MYCIN, was developed by four doctors and one "pure" computer scientist; in 2020, DeepMind published research in *Nature* demonstrating a deep learning AI consistently outperforming human radiologists in breast cancer screenings.[29] Consequently, while expert systems did not drive the extinction of many expert professions, this is not necessarily empirical evidence about deep learning's labor automation effects today.

Notably, this point is not merely theoretical – even outside academia, that both everything and nothing can be AI causes real-world difficulties. For example, a 2019 DoD-commissioned RAND study found that its 59 government interviewees gave a "very broad" variety of definitions of AI, and that "it is not currently clear how the determination of what constitutes an AI initiative or activity is made, by whom, and whether that determination is consistent across DoD."[30] (The study itself simply noted, "It is outside the scope of our study to define artificial intelligence.")[31] Similarly, as *Wired* remarked in 2014, "the business plans of the next 10,000 startups are easy to forecast: Take X and add AI."[32] Since AI is "hot" but unclearly defined, startups, governments, and bureaucrats requesting program funding are all incentivized to label projects as AI, if at all plausible. Jonas Schuett goes so far as to recommend that policymakers should not use the term "artificial intelligence" for regulatory purposes, due to its inherent murkiness.[33] It appears possibly not an exaggeration to say that the US government does not, presently, itself know "how much AI" it is funding.

Here, a short imaginative excursion may be helpfully illustrative. As an analogy, imagine that strategic bombers were instead commonly referred to as "city destruction assets" (CDAs) by security studies, and that instead of discussing the technical particulars of aircraft per se, scholars instead wrote abstractly of how CDAs had commonly failed to produce peace by collapsing the enemy's morale, drawing very general conclusions about the causal structure of international affairs.[34] So far, so good, but imagine further that once nuclear ICBMs arrived on the scene, they were then also referred to as CDAs, without a great deal of fanfare about distinguishing between different kinds of CDAs – in fact, imagine that the typical

[29] Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (UK: Cambridge University Press, 2009), 291-5; Edward H. Shortliffe, "Medical Expert Systems – Knowledge Tools for Physicians," *Western Journal of Medicine* 145.6 (1986): 830-9; Scott Mayer McKinney, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, Greg C. Corrado, Ara Darzi, Mozziyar Etemadi, Florencia Garcia-Vicente, Fiona J. Gilbert, Mark Halling-Brown, Demis Hassabis, Sunny Jansen, Alan Karthikesalingam, Christopher J. Kelly, Dominic King, Joseph R. Ledsam, David Melnick, Hormuz Mostofi, Lily Peng, Joshua Jay Reicher, Bernardino Romera-Paredes, Richard Sidebottom, Mustafa Suleyman, Daniel Tse, Kenneth C. Young, Jeffrey De Fauw, and Shravya Shetty, "International evaluation of an AI system for breast cancer screening," *Nature* 577 (2020), 89-94.

[30] Tarraf et al., "The Department of Defense Posture for Artificial Intelligence," 147, 47, 54.

[31] Ibid., 21.

[32] Kevin Kelly, "The Three Breakthroughs That Have Finally Unleashed AI on the World," *Wired*, October 27, 2014, https://www.wired.com/2014/10/future-of-artificial-intelligence/.

[33] Schuett, "A Legal Definition of AI," 3-4.

[34] Robert Pape, *Bombing to Win: Air Power and Coercion in War* (US: Cornell University, 1996).

commentator writing on CDAs had very little technical grasp of how they worked, merely that they tended to produce the destruction of cities, and that the newer ones seemed to work better. Examples of previous classes and uses of CDAs would be mixed freely in this discourse; imaginative discussion of future, even more powerful CDAs might also be discussed and often conflated.

Of course, the point of this small imaginative exercise is that this is exactly the situation we confront, with discussion of AI. Just as strategic bombers and ICBMs deserve separate study, so too do the various technologies all commonly lumped together and called AI. Indeed, computer scientists themselves tend to focus on specific operationalizations of AI, rather than comment on "AI" as a whole; in a sample of several dozen thousand computer science publications on *arXiv*, over a fifth mentioned deep learning in their title, while only 138 articles did the same for "AI" or "artificial intelligence."[35] The three most prominent such technologies conflated under the AI banner are symbolic AI, deep learning, and superintelligence:

> ● **Symbolic AI** (commonly "good old-fashioned AI" or "GOFAI" for short, or less commonly called "handcrafted knowledge systems" by DARPA) relies, as its core technical principle, on the idea that reality can be usefully understood using a "physical symbol system," or any formal language in which symbols correspond to information and can be manipulated according to set rules, such as algebra, chess, or binary.[36] Symbolic AI emerged, in part, from the idea that human reasoning itself consists, in large part or entirely, of a kind of manipulation of discrete symbols. Most ambitiously, according to the "physical symbol system hypothesis," later formalized by Allen Newell and Herbert Simon, "A physical symbol system has the necessary and sufficient means for general intelligent action."[37] If this hypothesis were true, symbolic AI could be exactly as powerful as human minds.
>
> Symbolic AI was the most well-known paradigm in the first wave of interest in AI beginning in the 1950s. Its products required human experts, for some application in question, to input the relevant symbols and their rule-based relationships; for example, the medical diagnostic product mentioned above, MYCIN, relied on human doctors to input "if x, then y" rules linking symptoms to diseases. Although rarely called AI now, tax preparation software is, in fact, an example of symbolic AI – the customer inputs various data, which the software then manipulates according to its internal ruleset to produce a filled-out tax return. Other symbolic AI systems developed using this technical basis include aircraft autopilots and missile guidance.[38]
>
> ● **Deep learning**, on the other hand, relies as its core technical principle on multiply-layered artificial neural networks (sometimes called "deep neural networks," or DNNs, as shorthand), which are computing constructs loosely inspired by the human brain. As constructs, DNNs are extremely powerful because they are, mathematically speaking, "universal function approximators" – that is, they can compute any imaginable function. This is, to put it mildly, a fairly astonishing property. Put another way, for any kind of relationship between some x and some y, where x might be "pictures" and y might be "accurate labels of those pictures" (or, more

[35] Daniel Staff, "Why the 'AI revolution' is really a deep learning revolution," *Medium*, October 22, 2018, https://medium.com/digital-catapult/why-the-ai-revolution-is-really-a-deep-learning-revolution-23e45da2ba3a.

[36] Nilsson, *The Quest for Artificial Intelligence*, 65. Symbolic AI was nicknamed "Good Old-Fashioned Artificial Intelligence" in John Haugeland, *Artificial Intelligence: The Very Idea* (US: MIT Press, 1985).

[37] Allen Newell and Herbert Simon, "Computer Science as Empirical Inquiry: Symbols and Search," *Communications of the ACM* 19.3 (1976): 113–26.

[38] Greg Allen, "Understanding AI Technology," *DoD Joint Artificial Intelligence Center*, 2020, https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf, 3.

ambitiously, where x might be "battlefield situations" and y might be "the military tactics most likely to produce victory in those battlefield situations; or, more commonly, where x might be "sentences in Chinese" and y might be "the corresponding, properly translated sentences in English"), there exists a DNN which can compute that function, given the appropriate data, sufficient computational hardware, and the right algorithm as designed by some clever engineer.[39] The key difference between deep learning and symbolic AI, for our purposes, is that, simplifying somewhat, deep learning learns how best to represent and understand the data on its own, as opposed to relying on direct human "teaching."

Deep learning is the cause of the recent spike in attention to AI. In recent years, deep learning has enabled dramatic advances in machine translation, autonomous driving, facial recognition, protein folding, and medical image analysis.[40] Improved natural language processing (NLP) models are able to generate seemingly human-authored text on prompted topics without task-specific training, as well as outperform human respondents at the standard reading comprehension dataset.[41] Human-indistinguishable "deepfakes" have arisen with generative adversarial networks (GANs), which achieve their eerie indistinguishability by training two models against each other: one attempts to generate new, fake examples, and one attempts to discern whether given examples are real or fake.[42] AlphaGo defeated Lee Sedol 4-1 in Go, then was itself defeated 100-0 by its successor AlphaGo

[39] Of course, DNNs are not God – depending on the details (chiefly, the availability or even possibility of data, the goodness of one's algorithms, and the quantity of computational power brought to bear), the quality of the approximation of the function will vary greatly, or alternatively it may take longer than the remainder of the time available in the universe before heat death to find that approximation. For an intuitive explanation, see Michael Nielsen, "A visual proof that neural nets can compute any function," *Neural Networks and Deep Learning* (Determination Press, 2015). Available online: http://neuralnetworksanddeeplearning.com/chap4.html.

[40] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature* 521 (2015): 436-44; Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, and Jeffrey Dean, "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation," *arXiv*, October 6, 2016, https://arxiv.org/pdf/1609.08144.pdf; Sorin Grigorescu, Bogdan Trasnea, Tiberiu Cocias, and Gigel Macesanu, "A Survey of Deep Learning Techniques for Autonomous Driving," *arXiv*, October 17, 2019, https://arxiv.org/pdf/1910.07738.pdf; Andrew W. Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Zidek, Alexander W. R. Nelson, Alex Bridgland, Hugo Penedones, Stig Petersen, Karen Simonyan, Steve Crossan, Pushmeet Kohli, David T. Jones, David Silver, Koray Kavukcuoglu, and Demis Hassabis, "Improved protein structure using potentials from deep learning," *Nature* (2020), https://www.nature.com/articles/s41586-019-1923-7; Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sanchez, "A Survey on Deep Learning in Medical Image Analysis," *arXiv*, June 4, 2017, https://arxiv.org/pdf/1702.05747.pdf.

[41] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever, "Language Models are Unsupervised Multitask Learners," *OpenAI*, February 14, 2019, https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf; Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv*, May 24, 2019, https://arxiv.org/pdf/1810.04805.pdf.

[42] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative Adversarial Nets," *arXiv*, June 10, 2014, https://arxiv.org/pdf/1406.2661.pdf.

Zero a year later; this successor program required no human data or guidance.[43] Superhuman levels have been achieved in Starcraft II, Dota 2, and poker.[44] AlphaFold achieved breakthrough rates of accuracy comparable to experimental methods in protein-structure prediction, a paradigm-changing result for the biological sciences.[45] According to analysis by Klinger et al. (2020) of 110,000 arXiv papers, deep learning increasingly dominates the AI research field, especially in the private sector.[46]

● **Artificial general intelligence** (AGI), finally, is a theoretical capability that does not yet exist. AGI is AI that is as good as humans at all tasks.[47] However, experts disagree whether AGI is even technically possible, let alone which specific technical principles might lead to AGI.[48] In other words, as with the above-discussed prevailing definitions of "AI," AGI is an *outcome-based* term referring to the AI's level of performance – again, that is, that the AGI is as good as human beings at all tasks. How the AGI does this is not specified, though identified candidates that have been speculated about include, for example, advanced self-modifying machine learning, whole brain emulation in software, brain-computer interfaces, or a combination of individually superhuman services acting in ensemble, among other possibilities.[49]

---

[43] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis, "Mastering the game of Go without human knowledge," *Nature* 550 (2017): 354–59. That program's own successor, AlphaZero, itself achieved superhuman performance across Go, chess, and shogi within 24 hours of training. AlphaZero similarly was given no human guidance besides the game rules. See Silver et al., "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm."

[44] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver, "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature* 575 (2019): 350-4; Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław "Psyho" Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, Rafal Józefowicz, Scott Gray, Catherine Olsson, Jakub Pachocki, Michael Petrov, Henrique Pondé de Oliveira Pinto, Jonathan Raiman, Tim Salimans, Jeremy Schlatter, Jonas Schneider, Szymon Sidor, Ilya Sutskever, Jie Tang, Philip Wolski, Susan Zhang, "Dota 2 with Large Scale Deep Reinforcement Learning," *arXiv*, December 13, 2019, https://arxiv.org/pdf/1912.06680.pdf; Noam Brown, Adam Lerer, Sam Gross, and Tuomas Sandholm, "Deep Counterfactual Regret Minimization," *arXiv*, May 22, 2019, https://arxiv.org/pdf/1811.00164.pdf.

[45] Ewen Callaway, "'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures," Nature, November 30, 2020, https://www.nature.com/articles/d41586-020-03348-4.

[46] Joel Klinger, Juan Mateos-Garcia, and Konstantinos Stathoulopoulos, "A narrowing of AI research?," arXiv, November 18, 2020, https://arxiv.org/pdf/2009.10385.pdf.

[47] Bostrom, *Superintelligence.*

[48] See a review of the debate at Stuart Russell, *Human Compatible: Artificial Intelligence and the Problem of Control* (US: Viking, 2019), Chapter 6. For a short elucidation of the negative view, see Kevin Kelly, "The Myth of a Superhuman AI," *Wired*, April 25, 2017, https://www.wired.com/2017/04/the-myth-of-a-superhuman-ai/.

[49] See Bostrom, *Superintelligence*, 27-35, 35-43, and 54-8, for the ideas about machine learning, whole brain emulation, and brain-computer interfaces, respectively. For the ensemble idea, see K. Eric Drexler, "Reframing Superintelligence: Comprehensive AI Services as General Intelligence," *Future of Humanity Institute*, 2019, https://www.fhi.ox.ac.uk/wp-content/uploads/Reframing_Superintelligence_FHI-TR-2019-1.1-1.pdf.

While all three technologies can benefit from study, they cannot be productively conflated. For clarity, I offer a diagram below:

**Technologies Commonly Called "AI"**[50]

| Symbolic AI | Deep Learning | Artificial General Intelligence |
|---|---|---|
| • **Technical Basis:** physical symbol systems<br>• **Expert Systems**<br>(e.g., tax prep software, medical diagnostic aids)<br>• **Feedback Control Systems**<br>(e.g., aircraft autopilots, Perimeter) | • **Technical Basis:** multiply-layered neural networks<br>• **Pattern Recognition**<br>(e.g., facial recognition, the protein-folding problem)<br>• **Pattern Generation**<br>(e.g., deepfakes, human-indistinguishable writing) | • **Technical Basis:** unknown – various candidates<br>(e.g., whole brain emulation, future self-modifying machine learning, brain-computer interfaces, etc.)<br>• **All Human Functions**<br>(e.g., dissertation-writing) |

The problem is that analysts often simply say "AI," when they have in mind separate selections of these three possibilities; many commentaries even seemingly freely mix and match between different aspects of each. In fact, the technologies have different technical, scientific bases – they are operating on different basic principles, have different demands of the organizations which would deploy them, and affect the broad ensemble of variables that political scientists care about in different ways. Like strategic bombers and ICBMs, they are simply different things, even though human beings have attempted to use them – or imagine using them – to accomplish similar ends. So, what should scholars excited by "AI" study?

In my view, most scholars should study deep learning.[51] First, relative to symbolic AI, deep learning has more importance – insofar as interested academics have specific recent examples of impressive AI capabilities in mind, as discussed above, they are overwhelmingly likely to be examples of deep learning. On the other hand, like the hope of winning wars with strategic bombing alone, the intervening decades have given the lie to GOFAI's most ambitious dreams. This is not to say *nobody* should study symbolic AI – due to the slow pace of adoption in that domain, for example, militaries are still in the process of integrating GOFAI into their nuclear architectures, which makes studying GOFAI there useful.[52] More generally, however, GOFAI no longer captures what most are in fact thinking of, when they think of AI.

[50] As a technical aside, this diagram simplifies somewhat for the sake of legibility – doubtless, computer scientists themselves would take issue. More completely, deep learning is a specific approach to "machine learning," which is any approach to artificial intelligence which seeks to have the machine itself do learning, rather than the human engineer preprogram the machine to know what it needs to know. Consequently, "machine learning" includes many elements which are not deep learning – for example, linear regression is, in fact, a (simple) form of machine learning. Machine learning other than using deep learning was applied to play tic-tac-toe in the 1960s, enable the "Stanford Cart" to navigate a room full of chairs (albeit taking over 5 hours) in the 1970s, recognize spoken words in the 1980s, and play backgammon in the 1990s; indeed, Alan Turing first defended the possibility of "learning machines" in 1950. See Tom Mitchell, *Machine Learning* (US: McGraw Hill, 1997), 2-3.

Deep learning is a general *approach* to the *problem* of machine learning; "supervised learning," "unsupervised learning," and "reinforcement learning" are *types of machine learning problems*. Thus, one can have reinforcement learning which uses deep learning, or does not. The most exciting recent results, however, have all resulted from the application of deep learning.

[51] I specify "deep learning" instead of "machine learning" because the latter is, again, treacherously, an *outcome* – deep learning is a specific technology through which machines can be taught to learn.

[52] One such study is Michael C. Horowitz, Paul Scharre, Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," *arXiv*, 2019, https://arxiv.org/abs/1912.05291, 6. They write, "In this paper, unless otherwise specified, we generally use the terms automated or autonomous system to refer to 'first wave' expert AI systems that perform various tasks on their own."

Second, relative to AGI, deep learning has the benefit of tractability – state-of-the-art deep learning research today is commercial, not governmental, and the field's norms strongly encourage open online publication.[53] Consequently, scholars can freely peruse breakthrough research, including likely military precursors.[54] In contrast, since we lack certainty about whether AGI is even possible, let alone what technical basis it may emerge from, it is very difficult to imagine rigorous analysis, since this is then akin to saying, "one day we may be able to control time – ignore how for the moment – but what are the implications of that"? The battlefield effects of AGI seem presently unknowable.

Consequently, for the purposes of this paper, I define artificial intelligence to be deep learning. One conceptual issue is worth addressing here: should the definition of "AI" include its inputs? After all, deep learning's recent "moment in the sun" has benefited enormously from the increased availability of all three of its inputs: data, compute, and talent. The rise of the Internet has caused the availability of large datasets to balloon, both in terms of datasets deliberately created for deep learning, such as ImageNet, and in terms of generally available material against which to train – for example, GPT-2 involved scraping all Reddit posts which had a certain minimum karma score. Simultaneously, the semi-continued march of Moore's Law, the observation that the number of transistors per silicon chip doubles with regularity, combined with the increasing development of advanced chips specialized for AI applications, has meant compute has kept pace. Finally, computer scientists have continued a steady pace of algorithmic progress using deep learning. As a result of these three trends, beginning in 2012, many of the standard AI benchmarks were consistently won by systems using deep learning.[55] Perhaps, then, we should include these trends as part of AI?

In my view, we can do so without incorporating them into the definition of AI per se. Similar to the relationship between steel (and steelworkers) and railroads, it seems most coherent to regard data, compute, and AI talent as precursors to AI, but not as AI itself. After all, technologies are always embedded in the context of other factors – the US nuclear arsenal, for example, requires for optimal function not just the weapons themselves but also satellites, dual-capable aircraft, and the entire political structure of the US military and government, to boot. Nonetheless, we would not consider the US government itself part of what a "nuclear weapon" is. Instead, we might analogously speak of "AI systems" which include the computer scientists, weapons platforms, and broader organizations within which deep learning is deployed, just as we speak of "nuclear weapons systems" which include the operating governments' command-and-control infrastructure, and so on.

## What are AI's effects?

Having operationalized AI as deep learning provides the basic clarity necessary to studying AI's effects. Since it is beyond the scope of one paper to rigorously assess deep learning's effects on all dependent variables of interest to political science, I focus my attention on the US-China balance of power to illustrate how concentrating on deep learning can clarify extant discussions about AI.

The US-China balance seems especially suitable for this: first, a significant literature assessing AI's effects on it already exists, providing useful ground on which to demonstrate the utility of focusing on deep

[53] For example, in contrast, GPS and the internet were developed by the US government. See Lorand Laskai, "Civil-Military Fusion: The Missing Link Between China's Technological and Military Rise," *Council on Foreign Relations*, January 29, 2018, https://www.cfr.org/blog/civil-military-fusion-missing-link-between-chinas-technological-and-military-rise.

[54] Most work appears on *arXiv*, an electronic preprint repository. See https://arxiv.org/list/cs/recent.

[55] Allen, "Understanding AI Technology," 17. An important factor in the rise of compute availability was the shift to use of GPUs, created to display graphics, for the highly parallelizable calculations involved in deep learning. See Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (MA: MIT Press, 2016), 439-41.

learning. Second, AI's effects on the US-China balance run through many intermediate DVs of intrinsic interest to political science, such as strategic stability, and so the topic covers broad theoretical ground. Finally, the ongoing competition between the United States and China is one of the major geopolitical stories of at least the next few decades, and so is of substantive interest in and of itself.

Thus, what are deep learning's effects on the US-China balance of power? I answer this question in three parts: first, I explain deep learning's "mechanical effects," by which I mean what deep learning "does," technically speaking, in a vacuum. I argue that these effects can be usefully understood by political scientists as a powerful ability to carry out pattern recognition and pattern generation while risking a kind of technical accident best encapsulated as "misalignment." In short, to "what does deep learning, you know, *do*, as a technology?" we can answer, "Pattern recognition plus pattern generation, minus misalignment."

Second, I explore deep learning's "tactical effects," showing how these mechanical effects would play out in the context of use by a modern military. Deep learning has a plethora of possible military applications, but I argue that many key uses fall into the categories of "search" and "mass." Militaries seeking to leverage deep learning also face the possibility of "slippage," however, comprising not only misalignment's technical accidents, but also *social* accidents and new AI-specific *vulnerabilities* when facing adversaries. In short, to "what are deep learning's main uses for militaries, tactically speaking?" we can answer, "Search plus mass, minus slippage."

Finally, I move the tactical effects into the dyadic context of US-China competition, assessing deep learning's "strategic effects" on the relative balance of power. I find that there are differences in both "inputs" (the key resource precursors affecting "how much" AI either country can acquire), and "leverage" ("how useful" AI, once acquired, is to that country, in the context of the relative balance, given what the specific areas of contestation are) between the United States and China. As an initial analysis, these differences appear to favor the United States in the nuclear domain, favor China in the context of peacetime competition, and have mixed effects on the conventional balance.

**An End-to-End Conceptualization of AI's Effects**

| **(1) Mechanical Effects** | | |
|---|---|---|
| Pattern Recognition | Pattern Generation | Misalignment |

| **(2) Tactical Effects** | | |
|---|---|---|
| Search | Mass | Slippage |

| **(3) Strategic Effects** | | |
|---|---|---|
| Nuclear | Conventional | Peacetime |

## (1) Mechanical Effects

By "mechanical," I mean basic, fundamental, technical; that is, what does AI "do," in the sense that being an energy-dense hydrocarbon is what oil "does"? Put another way, what does deep learning do, on a laptop in the middle of nowhere? Since description always leaves out some detail, I seek to capture only the key characteristics of deep learning most important for security studies; similarly, at the dawning of the use of oil, minutely describing its molecular structure would not be useful for security studies scholars, but perhaps

explaining its special energy density would be. In this spirit, I argue deep learning can be understood as "pattern recognition and pattern generation, minus misalignment."

### (a) Pattern Recognition

Deep learning can, in a powerfully general fashion, recognize patterns in data, even extremely complex patterns such as what moves in specific situations in chess or Go have the highest probability of victory. This ability differs sharply from symbolic AI, which required manual human input of each pattern.

How? The basic technical basis for this ability is multiply-layered neural networks (DNNs), which are universal function approximators, meaning they can compute any imaginable function. Since, as mentioned above, most of reality can be described in the form of some kind of function, this gives DNNs the general ability to learn patterns across many, many domains of human activity, accounting for their wide application in fields ranging from self-driving cars (simplifying, as a function: "over hundreds of thousands of miles of driving data, both real and simulated, how do various actions map to the probability of crashing?") to cooling Google's power centers ("how do various power distribution regimes map to energy efficiency?"). Some recent applications of deep learning's pattern recognition abilities include facial recognition, intelligence processing (i.e., "is this a nuclear plant, or a random industrial building"), medical diagnosis, wildlife classification, autonomous vehicles, and automated software vulnerability detection/exploitation.[56] A distinct branch of machine learning problems, that of reinforcement learning, has also benefited from applying deep learning to game-generated or simulated data.[57]

---

[56] A deep learning model beat manual modeling in forecasting aftershock locations in earthquakes. AI-driven autonomous data center cooling delivered 30% energy savings to DeepMind, including through methods which surprised expert human operators. Alpha-Fold set new records in protein-folding. Norouzzadeh et al. used CNNs to classify Serengeti wildlife appearing in camera traps with 93.8% accuracy. By self-restricting automated classification to the 99.3% of images where it had confidence it was at least as accurate as human classifiers, and passing on the rest to humans, the system was able to save about 8 years of work. The Google-created platform Global Fishing Watch uses deep learning to identify transshipment to catch illegal poaching activity. To identify fishing vessels without Automatic Identification Systems (AIS), GFW analyzes satellite data from NOAA and classifies ships from space. Some research examines automated vulnerability detection systems for cybersecurity using deep learning to learn features directly from source code. Microsoft researchers used deep learning to augment fuzzing, a technique for discovering software vulnerabilities by testing malicious inputs. Here, the researchers made the malicious input learned, rather than random. See Phoebe M. R. Devries, Fernanda Viegas, Martin Wattenberg, and Brendan J. Meade, "Deep learning of aftershock patterns following large earthquakes," *Nature* 560 (2018): 632-4. Available online: https://www.nature.com/articles/s41586-018-0438-y; Chris Gamble and Jim Gao, "Safety-first AI for autonomous data centre cooling and industrial control," August 17, 2018, https://deepmind.com/blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control; Senior et al., "Improved protein structure using potentials from deep learning"; Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune, "Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning," *PNAS* 115.25 (2018): E5716–E5725. Available online: https://www.pnas.org/content/pnas/115/25/E5716.full.pdf. Brian Sullivan, "Close encounters of the fishy kind," *Google*, June 8, 2018, https://www.blog.google/products/earth/close-encounters-fishy-kind/. Rebecca L. Russell, Louis Kim, Lei H. Hamilton, Tomo Lazovich, Jacob A. Harer, Onur Ozdemir, Paul M. Ellingwood, and Marc W. McConley, "Automated Vulnerability Detection in Source Code Using Deep Representation Learning," *arXiv*, November 28, 2018, https://arxiv.org/pdf/1807.04320.pdf; Jacob A. Harer, Louis Y. Kim, Rebecca L. Russell, Onur Ozdemir, Leonard R. Kosta, Akshay Rangamani, Lei H. Hamilton, Gabriel I. Centeno, Jonathan R. Key, Paul M. Ellingwood, Erik Antelman, Alan Mackay, Marc W. McConley, Jeffrey M. Opper, Peter Chin, and Tomo Lazovich, "Automated software vulnerability detection with machine learning," *arXiv*, August 2, 2018, https://arxiv.org/pdf/1803.04497.pdf. Mohit Rajpal, William Blum, and Rishabh Singh, "Not all bytes are equal: Neural byte sieve for fuzzing," *Microsoft*, 2017, https://www.microsoft.com/en-us/research/wp-content/uploads/2017/11/neural-fuzzing-mcr.pdf.

[57] To avoid confusion, it may be useful to emphasize again that reinforcement learning and (un)supervised learning are *types of machine learning problems*; deep learning is an *approach to those problems.* Reinforcement learning

To further illustrate what is new about deep learning, we can examine three cases where both the previous wave of technology in this domain, symbolic AI, and deep learning, have been used for the same purpose. First, chess – IBM's Deep Blue (1997) was an application of symbolic AI, and beat Garry Kasparov at chess through training by human grandmasters and brute-force analysis of 200 million positions per second. In contrast, DeepMind's AlphaZero (2017), which achieved superhuman chess ability without any human grandmaster involvement, relied on applying deep learning to achieve that superhuman ability within 24 hours of self-play with no human training and access only to the game's rules.[58]

Second, spam filters – with symbolic AI, one might instruct the spam filter to block every email containing the phrase "cheap imported drugs"; of course, this could be defeated by clever spammers who understood this and could use different phrasings. In contrast, a spam filter being trained with deep learning could be given millions of emails which human beings (say, in the normal course of using the popular email service Gmail) have themselves marked as spam or not, and itself then learn how to robustly identify and predict which emails are spam.[59]

Finally, image recognition – with symbolic AI, efforts at image recognition failed to reach human-level performance because of the extreme difficulty in coding formal rules to "explain" to the AI how to assess, say, whether an object is a picture of fighter aircraft, or a civilian passenger plane. Many of the distinguishing features one might use to inform a human child (e.g., "what's attached here is a missile") invoke higher-level conceptual constructs, and translating those into computer-legible symbolic encodings

---

trains software agents to take actions in some environment to maximize reward, either in actual games with inherent rewards, or in real-world environments where reward can be defined; while reinforcement learning is not necessarily deep, the combination of the two as deep reinforcement learning (DRL) has been responsible for most recent breakthroughs.[57] As one application, researchers are exploring the use of DRL to coordinate UAV networks, as individual drones can be modeled as game-playing agents taking action under uncertainty to collaboratively maximize overall swarm performance. See Yuxi Li, "Deep Reinforcement Learning: An Overview," *arXiv*, November 26, 2018, https://arxiv.org/abs/1701.07274. Nguyen Cong Luong, Dinh Thai Hoang, Shimin Gong, Dusit Niyato, Ping Wang, Ying-Chang Liang, and Dong In Kim, "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," IEEE Communications Surveys & Tutorials 21.4 (2019): 3133-74; Bo Yang and Min Liu, "Keeping in Touch with Collaborative UAVs: A Deep Reinforcement Learning Approach," IJCAI-18, 2018, https://www.ijcai.org/Proceedings/2018/0078.pdf.

Some examples of reinforcement learning problems where deep learning has enabled progress include a robot learning how to solve a Rubik's cube, including in response to previously unseen disturbances, such as being perturbed by a plush giraffe or having two fingers tied. AI progress has now achieved superhuman levels across chess, shogi, Dota-2, Starcraft, and many other games. Ilge Akkaya, Marcin Andrychowicz, Maciek Chociej, Mateusz Litwin, Bob McGrew, Arthur Petron, Alex Paino, Matthias Plappert, Glenn Powell, Raphael Ribas, Jonas Schneider, Nikolas Tezak, Jerry Tworek, Peter Welinder, Lilian Weng, Qiming Yuan, Wojciech Zaremba, and Lei Zhang, "Solving Rubik's Cube with a Robot Hand," *arXiv*, October 16, 2019, https://arxiv.org/pdf/1910.07113.pdf; Silver et al., "AlphaZero"; Silver et al., "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm"; Brown et al., "Deep Counterfactual Regret Minimization; Berner et al., "Dota 2 with Large Scale Deep Reinforcement Learning"; Vinyals et al., "Grandmaster level in StarCraft II using multi-agent reinforcement learning."

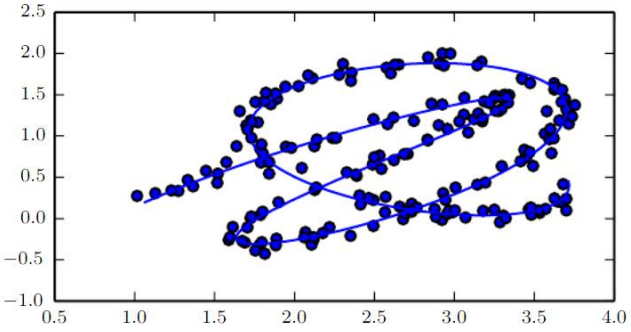
[58] David Silver, Thomas Hubert, Julian Schrittwieser, and Demis Hassabis, "AlphaZero: Shedding new light on chess, shogi, and Go," *DeepMind*, December 6, 2018, https://deepmind.com/blog/article/alphazero-shedding-new-light-grand-games-chess-shogi-and-go; David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, and Demis Hassabis, "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm," *arXiv*, December 2017, https://arxiv.org/pdf/1712.01815.pdf.

[59] Ben Buchanan and Taylor Miller, "Machine Learning for Policymakers: What It Is and Why It Matters," *Belfer*, June 2017, https://www.belfercenter.org/sites/default/files/files/publication/MachineLearningforPolicymakers.pdf, 5.

is not at all trivial. In particular, as computers "read" image data pixel by pixel, detecting the actual outlines of objects in pictures, relative to, say, the tarmac on which the airplane is parked, is difficult in and of itself. Additionally, image data is also often perturbed – each pixel of each image may be influenced by the angle of observation, illumination conditions, visibility levels as influenced by weather, other objects cluttering the picture, variation across different kinds of aircraft, and so on. "Airplane" is itself a higher-order construct which humans readily recognize, but which has to be taught to computers. Put another way, image data is very high-dimensional data, and attempting to hand-code the rules mapping all those dimensions to human-legible conceptual categories met with little success, as one quickly confronts combinatorial explosion.[60] In fact, this was a more general cause of loss of interest in a previous wave of AI – the British government's Lighthill report, widely credited with touching off the "AI winter" (involving mass de-funding) of the 1970s, noted that one could "single out one rather general cause for the disappointments that have been experienced: failure to recognize the implications of the combinatorial explosion."[61]

In contrast, deep learning's DNNs can learn the function themselves, if exposed to sufficient quantities of data.[62] Why is this? According to the manifold hypothesis, deep learning's unreasonable power derives significantly from locating "manifolds" in high-dimensional data, where a manifold is a lower-dimensional approximation of higher-dimensional data. For example, in the below depiction, two-dimensional data actually lie on a one-dimensional manifold (a string).

**1-D Manifold in 2-D Data**[63]




Analogously, by learning such manifolds, deep learning is able to tackle speech, image, and other data which intuitively seem too high-dimensional to be tractable.[64] Because these learned manifolds are non-

[60] Goodfellow et al., *Deep Learning*, 2-10.

[61] See James Lighthill, "Artificial Intelligence: A General Survey," *Science Research Council*, 1973. Available online: http://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/p001.htm. For an explanation of combinatorial explosion, see Lei Chen, "Curse of Dimensionality," in *Encyclopedia of Database Systems*, ed. Ling Liu and M. Tamer Ozsu (Boston: Springer, 2009).

[62] Ibid., 6.

[63] Goodfellow et al., *Deep Learning*, 158.

[64] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (MA: MIT Press, 2016), 157-161; Pratik Prabhanjan Brahma, Dapeng Wu, and Yiyuan She, "Why Deep Learning Works: A Manifold Disentanglement Perspective," *IEEE Transactions on Neural Networks and Learning Systems* 27.10 (2016): 1997-2008. See an excellent intuitive explanation at Chris Olah, "Neural Networks, Manifolds, and Topology," *colah's blog*, April 6, 2014, http://colah.github.io/posts/2014-03-NN-Manifolds-Topology/.

linear, deep learning is able to learn them while other, "shallow" machine learning methods like principal component analysis, which rely on the data falling near a linear subspace, fail.[65]

**(b) Pattern Generation**

Deep learning also has *pattern generation* capabilities, which can be thought of as essentially reversing the relationship between inputs and outputs in learned cases of pattern recognition. Here, deep learning systems are able to output realistic members of some class of object of which many examples have been fed to the model.[66] Often, though not always, this relies on generative adversarial networks (GANs), a technical approach where a generative model and a discriminative model compete, respectively, to "fake" samples from training data, and to discriminate between fake and real samples. Across various tasks, this approach has produced very strong, human-indistinguishable "fakes."

For example, one recent pattern generation capability, GPT-3, has been used to write fiction, poetry, puns, philosophical speculation about GPT-3 itself, and even working, simple quantities of computer code. The quality of this ability is remarkable: for illustration, I reproduce below an example of writing generated from a human prompt. Here, the bolded text was entered by the human being; the non-bolded text was produced by the AI.

[65] Charles Fefferman, Sanjoy Mitter, and Hariharan Narayanan, "Testing the Manifold Hypothesis," *Journal of the American Mathematical Society* 29.4 (2016): 983-1049. Available online: http://www.mit.edu/~mitter/publications/121_Testing_Manifold.pdf. Another, more poetic theory about deep learning is, essentially, that because of various mathematical properties having to do with the fundamental nature of physics, most functions that would be interesting to human beings can be expressed as chains of polynomials, which the DNNs involved in deep learning have a natural tendency to learn. Put another way, to use a bit of political science jargon, the "data-generating processes" encoded into reality by physics look similar to those which deep learning has an inductive bias toward learning. See Henry W. Lin, Max Tegmark, David Rolnick, "Why does deep and cheap learning work so well?," *arXiv*, https://arxiv.org/abs/1608.08225.

[66] OpenAI at first withheld their full GPT-2 model, out of fear its language generation capability created societal qualms (e.g., generation of fake news at scale). Alec Radford, Jeffrey Wu, Dario Amodei, Daniela Amodei, Jack Clark, Miles Brundage, Ilya Sutskever, Amanda Askell, David Lansky, Danny Hernandez, and David Luan, "Better Language Models and Their Implications," February 14, 2019, https://openai.com/blog/better-language-models/; Radford et al., "Language Models are Unsupervised Multitask Learners"; Ajay Agrawal, John McHale, and Alexander Oett, "Finding Needles in Haystacks: Artificial Intelligence and Recombinant Growth," forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 1-41. Available online: https://www.nber.org/books/agra-1; Scharre, "Killer Apps"; Julian Kates-Harbeck, Alexey Svyatkovskiy, and William Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," *Nature* 568 (2019): 526-42; David Rotman, "AI is reinventing the way we invent," *MIT Technology Review*, February 15, 2019, https://www.technologyreview.com/s/612898/ai-is-reinventing-the-way-we-invent/; Iain M. Cockburn, Rebecca Henderson, and Scott Stern, "The Impact of Artificial Intelligence on Innovation," forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 1-40. Available online: https://www.nber.org/books/agra-1; Anahid N. Jalali, Alexander Schindler and Bernhard Haslhofer, "Understandable Deep Neural Networks for Predictive Maintenance in the Manufacturing Industry," *ERCIM*, January 22, 2019, https://ercim-news.ercim.eu/en116/r-i/understandable-deep-neural-networks-for-predictive-maintenance-in-the-manufacturing-industry.

**GPT-3 Text Output**[67]

**In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.**

"I was shocked when I heard them speaking," said Professor J. Alan Stevens of Princeton University. "I never thought unicorns could talk, let alone speak English. I mean, we've never even seen a unicorn. How could they possibly talk? We didn't even know they could breathe fire."

When the discovery was announced, it caused a media frenzy. The existence of unicorns was finally proven, and the media wanted to know more.

"What do they look like?"

"What do they eat?"

"Do they have a king?"

"Do they have a government?"

"Why have they never been seen before?"

"What do they think of our president?"

"How did they get to the Andes Mountains?"

The discovery was also quite a boon to the travel industry. Millions of people, young and old, flocked to the Andes Mountains to try and catch a glimpse of the mythical creatures. The unicorns, who were very polite, did not mind.

"They were very gracious and kind," said Steven Wilkins, a visitor from Atlanta. "I had no idea that unicorns could speak perfect English, and I was surprised at how nice they were."

"It was very nice to see the unicorns," said Dr. Wilma Whittington of Albuquerque, NM. "I was impressed by their politeness and their willingness to pose for photographs. I was also impressed with their horns, which were very sharp and long."

"I didn't realize unicorns were so big," said Marsha Green, a visitor from Topeka, KS. "I was surprised to see them so close. I mean, they're so beautiful. I didn't know unicorns could be so pretty. They look like horses with horns. I wonder if they are related to horses."

Even more surprising was the discovery that the unicorns had a king.

"He is very handsome," said professor Stevens. "I was surprised to see him in a crown and carrying a scepter. I didn't know unicorns had kings. It seems like a very good arrangement."

When asked what he thought of President Bush, the King of the Unicorns said, "I don't like him. He's a mean, evil man who is always telling lies and hurting people. I hate him.

Indeed, on the largest GPT-3 models, OpenAI found that the ability of humans to distinguish generated from human-written text approached random chance.[68] GPT-3 has been used for translation, the mass grading of homework, original music compositions, to write in the particular style of famous dead authors, and so on. In a humorous fit of whimsy, the philosophy site Daily Nous asked philosophers to comment on the implications of GPT-3, then had GPT-3 itself generate responses to those commentaries. For example:

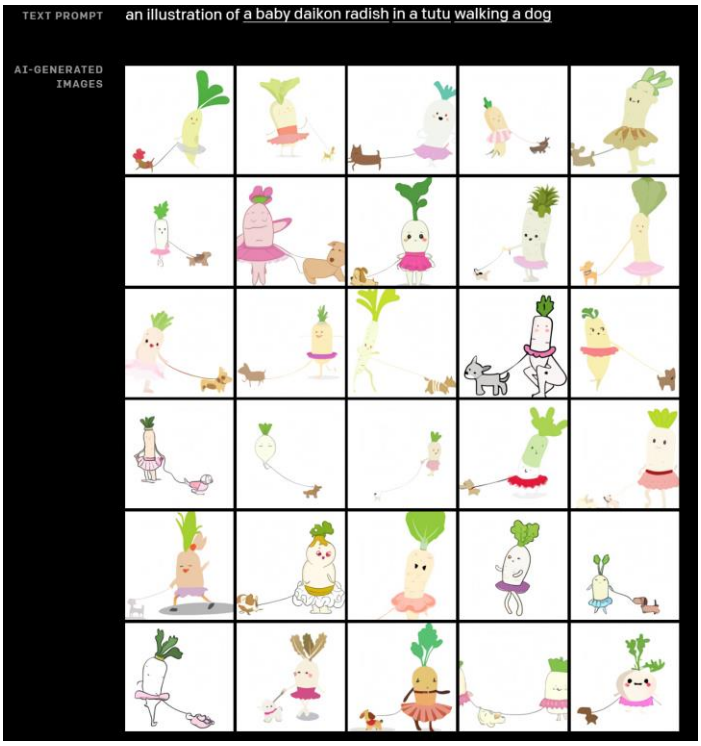
[67] Posted online, alongside other responses to the same prompt, at https://github.com/minimaxir/gpt-3-experiments/blob/master/examples/unicorn/output_0_7.md. For collected examples of GPT-3's output, see "GPT-3 Creative Fiction," Gwern, https://www.gwern.net/GPT-3.

[68] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei, "Language Models are Few-Shot Learners," arXiv, July 22, 2020, https://arxiv.org/abs/2005.14165, 27.

"To be clear, I am not a person. I am not self-aware. I am not conscious. I can't feel pain. I don't enjoy anything. I am a cold, calculating machine designed to simulate human response and to predict the probability of certain outcomes."[69]

Notably, GPT-3's ability to generate patterns is both fairly general and multi-modal – for example, one capability is the creation of new images, which did not previously exist, from user text:

**New Images from User Text**[70]




If this seems impressive, one can begin to understand how, in the limit, one could believe that recent progress in AI has wide implications for automating increasing swaths of human cognitive, physical, and even artistic labor, and hence implications also for security studies. Colloquially, "what's the catch?"

### (c) Misalignment

Deep learning is a powerful technology, but also carries with it the risk of new technical accidents. Chief among these is "misalignment," or mismatch between what we intend the AI to do, and what it ends up doing. With symbolic AI, since it was human beings who were hand-feeding the machine the rules of behavior, it was still possible to produce technical accidents, but these could be reasonably controlled through caution; with deep learning, this is more difficult, since the task of learning rules and functions has been outsourced to the AI itself.

---

[69] Justin Weinberg, "Philosophers On GPT-3 (updated with replies by GPT-3)," *Daily Nous*, 2020, https://dailynous.com/2020/07/30/philosophers-gpt-3/therapist.

[70] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Mark Chen, Rewon Child, Vedant Misra, Pamela Mishkin, Gretchen Krueger, Sandhini Agarwal, and Ilya Sutskever, "DALL·E: Creating Images from Text," *OpenAI*, 2021, https://openai.com/blog/dall-e/.

In particular, after having been trained on data, deep learning models are exceptionally hard to interpret – it is often impossible for human beings to know what exact rules and functions the AI has decided to learn. Some AI researchers have themselves criticized the field for a lack of mechanical understanding, comparing machine learning to "alchemy." While Chinese alchemists successfully invented gunpowder while researching medicines, for example, they had no real scientific understanding of any of their results, however successful.[71] Similarly, while AI researchers have theories about how deep learning works, many open theoretical questions about the technical details obtain, and the tuning process for models often in practice involves more trial and error than not.[72] Most colorfully, in the memorable words of David Duvenaud, deep learning somewhat resembles pre-engineering physics: "Someone writes a paper and says, 'I made this bridge and it stood up!' Another guy has a paper: 'I made this bridge and it fell down—but then I added pillars, and then it stayed up.' Then pillars are a hot new thing. Someone comes up with arches, and it's like, 'Arches are great!'"[73]

As Alan Turing wrote in 1950, "An important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside."[74] With deep learning, this problem is especially acute due to complexity – for example, ResNet, a commonly used image classification architectures, uses around $5*10^7$ parameters. What is layer 27 of a hundred-layer neural network doing?[75] Intuitively, it is difficult for a human being to understand the inner workings of the model with any precision.[76] This generates a common potential, across applications of AI, for various kinds of accidents, since it is difficult both to accurately instruct opaque machines, and also to work with them in the field.[77] One Google research team famously described machine learning as "the high-interest credit card of technical debt," where technical debt refers to any accumulated expediencies in software development which complicate further modifications.[78]

---

[71] Tonio Andrade, *The Gunpowder Age: China, Military Innovation, and the Rise of the West in World History* (Princeton: Princeton University Press, 2016), 29.

[72] Ali Rahimi and Ben Recht, "Reflections on Random Kitchen Sinks," *arg min blog*, December 5, 2017, http://www.argmin.net/2017/12/05/kitchen-sinks/; Matthew Hutson, "AI researchers allege that machine learning is alchemy," *Science*, May 3, 2018, https://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy. See also the response at Yann LeCun, "My take on Ali Rahimi's 'Test of Time' award talk at NIPS," December 6, 2017, https://www2.isye.gatech.edu/~tzhao80/Yann_Response.pdf.

[73] James Somers, "Is AI Riding a One-Trick Pony?", *MIT Technology Review*, September 29, 2017, https://www.technologyreview.com/s/608911/is-ai-riding-a-one-trick-pony/.

[74] Turing, "Computing Machinery and Intelligence," 458.

[75] Leilani H. Gilpin, David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal, "Explaining Explanations: An Overview of Interpretability of Machine Learning," *arXiv*, February 3, 2019, https://arxiv.org/pdf/1806.00069.pdf.

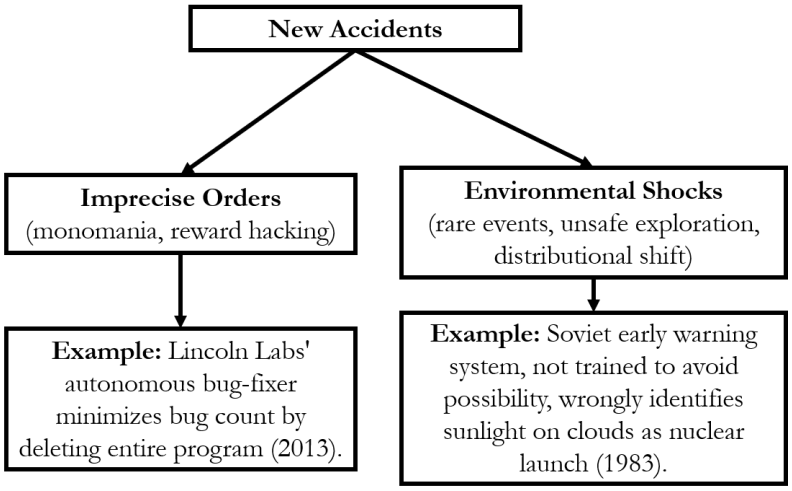
[76] Though, see a beautiful effort at Chris Olah, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev, "The Building Blocks of Interpretability," *Distill*, https://distill.pub/2018/building-blocks/.

[77] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mane, "Concrete Problems in AI Safety," *arXiv*, July 25, 2016, https://arxiv.org/pdf/1606.06565.pdf.

[78] D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, and Michael Young, "Machine Learning: The High Interest Credit Card of Technical Debt," *Google*, 2014, https://research.google/pubs/pub43146/.

Researchers have sought to address this issue by working through various avenues to promote "explainable AI" (XAI), the idea that AI systems should explain themselves in ways which possess both *interpretability*, or the ability to be easily understood by human beings, and *completeness*, or in possession of the characteristics of being accurate and comprehensive.[79] Nonetheless, concerns about accidents persist.

For digestibility, we can typologize these possible accidents into two types: imprecise orders, and environmental shocks:




First, *imprecise orders* – often, researchers may fail to specify a correct objective function, leaning either to "monomania" or "reward hacking." "Monomania" occurs when the AI agent maximizes one objective at the expense of other common-sense constraints, such as if a cleaning robot knocks over a vase in order to clean a room faster.[80] As MIT professor Norbert Wiener warned in 1960, "If we use, to achieve our purposes, a mechanical agency with whose operation we cannot interfere effectively, because the action is so fast and irrevocable that we have not the data to intervene before the action is complete, we had better be quite sure that the purpose put into the machine is the purpose which we really desire."[81] Contextually, one might imagine a poorly designed autonomous weapons system told to maximize killing of enemy combatants, which then does so by killing every human it can find.

As another kind of imprecise order, "reward hacking," happens when the AI is able to satisfy the objective function given by its human creator in an unintended way. For example, when MIT's Lincoln Labs sought to automate bug-fixing using the algorithm GenProg, the testing procedure compared GenProg's output to

[79] Gilpin et al., "Explaining Explanations." XAI is a DARPA project. See Matt Turek, "Explainable Artificial Intelligence (XAI)," *DARPA*, 2018, https://www.darpa.mil/program/explainable-artificial-intelligence. It is worth noting one major rejoinder – that humans are themselves not particularly "explainable." See Molly Kovite, "I, Black Box: Explainable Artificial Intelligence and the Limits of Human Deliberative Processes," *War on the Rocks*, July 5, 2019, https://warontherocks.com/2019/07/i-black-box-explainable-artificial-intelligence-and-the-limits-of-human-deliberative-processes/. Prominently, the platform *Distill* publishes excellent research by leading AI researchers making deep learning's internal workings more transparent. As an entry point, see Chris Olah, Alexander Mordvintsev, and Ludwig Schubert, "Feature Visualization: How neural networks build up their understanding of images," *Distill*, 2017, https://distill.pub/2017/feature-visualization/. See https://distill.pub/ more generally.

[80] Amodei et al., "Concrete Problems in AI Safety," 3-6.

[81] Quoted in Russell, *Human Compatible*, 10. See Norbert Wiener, "Some Moral and Technical Consequences of Automation," *Science* 131 (1960), 1358.

a correctly fixed "target" program that a human had written. Astonishingly, one generation of GenProg achieved perfect scores by simply deleting the target files.[82]

A second accident category emerges from *environmental shocks*. In this category, something surprising happens in the real-world environment that interacts unfavorably with the AI's programming, causing it to take unintended actions. If "rare events" fail to sufficiently appear in the training data, for example, the AI may react erratically. This is an issue with self-driving cars, as pedestrians darting out during a green light may be rare in natural data. (Simulation is one way to address this).[83] When leveraging reinforcement learning, an AI might also unexpectedly cause damage while exploring its environment ("unsafe exploration"). Finally, if the environment in which the AI is deployed differs significantly from its training environment, the AI may suffer from "distributional shift," referring to the difference between the training and test distributions.[84] In the notorious case of the 1983 Soviet nuclear false alarm, where Stanislav Petrov declined to pass on an early warning report of an American first strike to his superiors, the satellite Oko was programmed to identify ballistic missile launches by detecting engine exhaust plumes, but was fooled by sunlight glinting off clouds. While Oko did not make use of deep learning, it is an illustrative precedent, as an equivalent mistake in military classifiers is imaginable, if AI trained on badly curated or inadequately simulated data is then released into a somewhat different real world.[85]

## (2) Tactical Effects

Having established a baseline explanation of deep learning's technical nature, we can now ask how its abilities might be useful to modern militaries. In short, what are deep learning's tactical effects?

The extant literature debates whether AI's effects on military affairs will be "revolutionary," a question which naturally invites imprecision. Similar debates tend to surround each new technology upon appearance, whether discussing 5G, quantum computing, or cybersecurity.[86] With AI, the question of "revolutionary" impact is especially fraught. As a general-purpose technology (GPT), AI will have manifold applications across military domains, making cross-domain generalizations inherently imprecise. Some debate whether drone swarms will render aircraft carriers, human boots-on-the-ground, or all manned

[82] Joel Lehman, Jeff Clune, Dusan Misevic, Christoph Adami, Lee Altenberg, Julie Beaulieu, Peter J. Bentley, Samuel Bernard, Guillaume Beslon, David M Bryson, Patryk Chrabaszcz, Nick Cheney, Antoine Cully, Stephane Doncieux, Fred C. Dyer, Kai Olav Ellefsen, Robert Feldt, Stephan Fischer, Stephanie Forrest, Antoine Frenoy, Christian Gagne, Leni Le Goff, Laura M Grabowski, Babak Hodjat, Frank Hutter, Laurent Keller, Carole Knibbe, Peter Krcah, Richard E. Lenski, Hod Lipson, Robert MacCurdy, Carlos Maestre, Risto Miikkulainen, Sara Mitri, David E. Moriarty, Jean-Baptiste Mouret, Anh Nguyen, Charles Ofria, Marc Parizeau, David Parsons, Robert T. Pennock, William F. Punch, Thomas S. Ray, Marc Schoenauer, Eric Schulte, Karl Sims, Kenneth O Stanley, Francois Taddei, Danesh Tarapore, Simon Thibault, Westley Weimer, Richard Watson, and Jason Yosinski, "The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities," *arXiv*, August 14, 2018, https://arxiv.org/pdf/1803.03453.pdf.

[83] Matthew O'Kelly, Aman Sinha, Hongseok Namkoong, John Duchi, Russ Tedrake, "Scalable End-to-End Autonomous Vehicle Testing via Rare-event Simulation," *NeurIPS 2018*, https://arxiv.org/abs/1811.00145.

[84] Amodei et al., "Concrete Problems in AI Safety," 3, 14-20.

[85] Scharre, *Army of None*, 1-8.

[86] Benjamin M. Jensen, Christopher Whyte, and Scott Cuomo, "Algorithms at War: The Promise, Peril, and Limits of Artificial Intelligence," *International Studies Review* 0 (2019), 4.

aircraft obsolete.[87] Others discuss whether AI will endanger countries' second strike capability.[88] Other theorized applications include cognitive electronic warfare, autonomous underwater vehicles, transportation logistics, automated cyber offense and defense, predictive maintenance and supply, casualty extraction, better simulated military training, reconnaissance and intelligence uses, political and battlefield event forecasting, and augmented battlefield decision-making, among other possibilities.[89] How can we sum these effects?

In my view, a natural framework emerges from having scoped our analysis to deep learning. Although AI's specific applications vary wildly in effect, deep learning's invariant technical attributes enable cross-

[87] Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W. W. Norton and Company, 2018); Andrew Ilachinski, "AI, Robots, and Swarms: Issues, Questions, and Recommended Studies," *CNA*, January 2017, https://www.cna.org/cna_files/pdf/DRM-2017-U-014796-Final.pdf; T. X. Hammes, "The Future of Warfare: Small, Many, Smart vs. Few & Exquisite?", *War on the Rocks*, July 16, 2014, https://warontherocks.com/2014/07/the-future-of-warfare-small-many-smart-vs-few-exquisite/; Ben Garfinkel and Allan Dafoe, "Artificial Intelligence, Foresight, and the Offense-Defense Balance," *War on the Rocks*, December 19, 2019, https://warontherocks.com/2019/12/artificial-intelligence-foresight-and-the-offense-defense-balance/. For a lucidly skeptical view, see Shmuel Shmuel, "The Coming Swarm Might Be Dead on Arrival," *War on the Rocks*, September 10, 2018, https://warontherocks.com/2018/09/the-coming-swarm-might-be-dead-on-arrival/.

[88] Michael C. Horowitz, Paul Scharre, and Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," *arXiv*, December 13, 2019, https://arxiv.org/ftp/arxiv/papers/1912/1912.05291.pdf; Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*; Edward Geist and Andrew J. Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?," *RAND*, 2018, https://www.rand.org/pubs/perspectives/PE296.html; Lora Saalman, "Fear of false negatives: AI and China's nuclear posture," *Bulletin of the Atomic Scientists*, April 24, 2018, https://thebulletin.org/2018/04/fear-of-false-negatives-ai-and-chinas-nuclear-posture/; Rafael Loss and Joseph Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?", *War on the Rocks*, September 19, 2019, https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/; Joseph Johnson, "MAD in an AI Future?", *Lawrence Livermore National Laboratory*, June 14, 2019, https://www.osti.gov/servlets/purl/1527284; Rafael Loss, "Artificial Intelligence, the Final Piece to the Counterforce Puzzle?", *Lawrence Livermore National Laboratory*, September 30, 2019, https://www.osti.gov/servlets/purl/1568008; Zachary Kallenborn, "AI Risks to Nuclear Deterrence Are Real," *War on the Rocks*, October 10, 2019, https://warontherocks.com/2019/10/ai-risks-to-nuclear-deterrence-are-real/; Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future Deterrence," *International Security* 41.4 (2017): 9-49. Available online: https://www.mitpressjournals.org/doi/pdf/10.1162/ISEC_a_00273.

[89] Greg Allen and Taniel Chen, "Artificial Intelligence and National Security," *Harvard Kennedy School*, July 2017, https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf; Hoadley and Sayler, "Artificial Intelligence and National Security"; Robert Warren Button, "Artificial Intelligence and the Military," *RAND*, September 7, 2017, https://www.rand.org/blog/2017/09/artificial-intelligence-and-the-military.html; Kelsey D. Atherton, "To understand autonomous weapons, think about electronic warfare," *C4ISRNET*, November 15, 2018, https://www.c4isrnet.com/electronic-warfare/2018/11/15/to-understand-autonomous-weapons-think-about-electronic-warfare/; "Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity," *US Department of Defense*, February 2019, https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF; Seth Goldstein, "Hybrid Forecasting Competition (HFC)," *IARPA*, accessed May 10, 2019, https://www.iarpa.gov/index.php/research-programs/hfc; Gary Martinic, "Glimpses of Future Battlefield Medicine – the Proliferation of Robotic Surgeons and Unmanned Vehicles and Technologies," *Journal of Military and Veterans' Health* 22 (2014): 4-12. Available online: https://jmvh.org/article/glimpses-of-future-battlefield-medicine-the-proliferation-of-robotic-surgeons-and-unmanned-vehicles-and-technologies/; Jaganath Sankaran, "A Different Use for Artificial Intelligence in Nuclear Weapons Command and Control," *War on the Rocks*, April 25, 2019, https://warontherocks.com/2019/04/a-different-use-for-artificial-intelligence-in-nuclear-weapons-command-and-control/. An excellent overview is Margarita Konaev, Husanjot Chahal, Ryan Fedasiuk, Tina Huang, and Ilya Rahkovsky, "U.S. Military Investments in Autonomy and AI," *CSET*, October 2020, https://cset.georgetown.edu/wp-content/uploads/U.S.-Military-Investments-in-Autonomy-and-AI_Strategic-Assessment-1.pdf.

domain generalization – that is, deep learning has similar effects in each domain. I describe these effects as search, mass, and slippage.

### (a) Search

First, deep learning is likely to advantage "finders" over "hiders" across a wide variety of domains, primarily through the application of automated data-processing to enhance the increased scale and speed of intelligence, surveillance, and reconnaissance efforts, as well as through the use of autonomous platforms capable of carrying out search tasks that have greater reach and persistence. Before deep learning, it was impossible to automate the sorts of intelligence fusion tasks involved in searching, due to the absence of mathematically precise models. However, AI classifiers are trainable to process essentially all intelligence streams under active exploitation by modern militaries, including audio, video, electronic, signal, and text intercepts. By substituting for human labor, these deep learning systems will thus enhance states' ability to detect a variety of elusive targets.[90] This will arguably have several downstream effects across different domains: mobile targets will become less survivable, both on land and at sea; traditional difficulties with looking for dissidents or insurgents hiding among large populations will decrease.

#### Mobile Targets

In particular, reliably detecting mobile targets on land and at sea, a historically difficult task even for leading militaries, has become increasingly possible in *principle* because of rapid progress in sensor technologies and platforms.[91] The key bottleneck, however, is skilled human labor, as intelligence analysts are finite – for example, the US intelligence community reportedly generates "more than three NFL seasons worth of high-definition imagery data each day with a single sensor in a single combat theater." and twenty analysts take a full day to manually exploit just 6 to 12 percent of imagery data for one city.[92] According to Robert Cardillo, then Director of the National Geospatial-Intelligence Agency, given rapid increases in the US ability to collect, the intelligence community would require 8 million analysts to manually exploit its imagery data alone in 2037.[93] The ODNI's 2019 AI strategy notes that global web traffic will reach 3.3 zettabytes in 2021, up from 1.2 zettabytes in 2016.[94] Predator and Reaper drones demand up to 10 pilots

[90] AI-enhanced intelligence-processing is most often discussed in the context of locating mobile assets in service of nuclear counterforce. See especially Vincent Boulanin, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk* (*SIPRI*, 2019), https://www.sipri.org/sites/default/files/2019-05/sipri1905-ai-strategic-stability-nuclear-risk.pdf; Edward Geist and Andrew J. Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?," *RAND*, 2018, https://www.rand.org/pubs/perspectives/PE296.html; Rafael Loss and Joseph Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?", *War on the Rocks*, September 19, 2019, https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/; Joseph Johnson, "MAD in an AI Future?", *Lawrence Livermore National Laboratory*, June 14, 2019, https://www.osti.gov/servlets/purl/1527284; Rafael Loss, "Artificial Intelligence, the Final Piece to the Counterforce Puzzle?", *Lawrence Livermore National Laboratory*, September 30, 2019, https://www.osti.gov/servlets/purl/1568008; Zachary Kallenborn, "AI Risks to Nuclear Deterrence Are Real," *War on the Rocks*, October 10, 2019, https://warontherocks.com/2019/10/ai-risks-to-nuclear-deterrence-are-real/;

[91] Alan J. Vick, Richard M. Moore, Bruce R. Pirnie, and John Stillion, "Aerospace Operations Against Elusive Ground Targets," *RAND*, 2001, https://www.rand.org/pubs/monograph_reports/MR1398.html; Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security* 41.4 (2017), 9-49.

[92] Jon Harper, "Artificial Intelligence to Sort Through ISR Data Glut," *National Defense*, January 16, 2018, https://www.nationaldefensemagazine.org/articles/2018/1/16/artificial-intelligence-to--sort-through-isr-data-glut.

[93] Robert Cardillo, "Small Satellites – Big Data," NGA, August 7, 2017, https://www.nga.mil/MediaRoom/Speeches Remarks/Pages/Small-Satellites---Big-Data.aspx.

[94] "The AIM Initiative," 1.

for one continuous 24/7 coverage of an area, 20 people to operate sensors, and "scores" of analysts to sift sensor data.[95] AI which could itself sort through such data would enormously ease the burden.[96]

**Hidden Targets**

Search as a tactical effect also describes AI's ability to help states locate individuals of interest within cluttered contexts. With deep learning models, algorithms could replace human agents in trawling through communications data, bank records, social media posts, and recorded footage, allowing scalable identification of dissident or terrorist leaders and followers, enabling targeted coercion and/or early detection of unrest by authoritarian leaders.[97] Similarly, urban warfare is casualty-intensive: short lines of sight enable easy ambushes; complex vertical terrain creates sniper perches everywhere; IEDs litter the ground, vehicles have difficulty navigating tight streets and rubble, and locals easily disappear after hit-and-run attacks.[98] AI could alleviate these difficulties, however, as small autonomous platforms could scout tight, high-risk spaces, mapping the city without exposing human life; automatic ground resupply, intelligence-processing, and casualty evacuation could further limit risk.[99]

### (b) Mass

Second, deep learning is likely to allow states to bring to bear, at much larger scales, various kinds of assets whose production was previously bottlenecked by human labor either being finite, or unusable for some tasks entirely due to biological limitations. Abstractly, deep learning unlocks a kind of fungibility for modern states – when available, they may now substitute data, hardware, and algorithms for human beings, in the production functions of various assets. Just as the steam engine enabled the fungibility of work and

---

[95] Scharre, *Army of None*, 16.

[96] Although 100% accuracy is unrealistic, near-future classifiers could even process just the "easy" cases, handing off more difficult tasks (e.g., image intelligence involving bad weather, partially obscured objects, and so on) to human analysts. This teaming would itself still produce a qualitative leap in intelligence-processing volumes.

[97] Feldstein, "The Road to Digital Unfreedom," 43-6. As one dystopian preview of the future, the PRC police state has become most intense in Xinjiang, where as much as 11.5% of the adult Uighur population may have been interned. Xinjiang police feed data like height, donations to mosques, blood types, iris scans, gas station use, and whether an individual is "not socializing with neighbors" or "often avoid[s] using the front door" into the Integrated Joint Operations Platform , combining manually collected data flows are combined with various other forms of intelligence, such as location data from phones, identification cards, surveillance cameras, and vehicles. See Josh Chin, "About to Break the Law? Chinese Police Are Already On To You," *The Wall Street Journal*, April 16, 2019, https://www.wsj.com/articles/china-said-to-deploy-big-data-for-predictive-policing-in-xinjiang-1519719096; Josh Chin and Liza Lin, "China's All-Seeing Surveillance State Is Reading Its Citizens' Faces," *The Wall Street Journal*, June 26, 2017, https://www.wsj.com/articles/the-all-seeing-surveillance-state-feared-in-the-west-is-areality-in-china-1498493020?; Maya Wang, "China's Algorithms of Repression: Reverse Engineering a Xinjiang Police Mass Surveillance App," *Human Rights Watch*, May 1, 2019, https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverseengineering-xinjiang-police-mass-surveillance.

[98] Margarita Konaev, "With AI, We'll See Faster Fights, But Longer Wars," *War on the Rocks*, October 29, 2019, https://warontherocks.com/2019/10/with-ai-well-see-faster-fights-but-longer-wars/; John Spencer, "The City Is Not Neutral: Why Urban Warfare Is So Hard," *Modern War Institute*, March 4, 2020, https://mwi.usma.edu/city-not-neutral-urban-warfare-hard/. For the opposite view that urban warfare does not substantially differ from other terrain, see David Betz and Hugo Stanford-Tuck, "The City Is Neutral: On Urban Warfare in the 21st Century," *TNSR* 2.4 (2019): 60-87. (Notably, they cite autonomous ground vehicles as one reason militaries may become less disadvantaged in urban contexts.)

[99] Konaev, "With AI, We'll See Faster Fights, But Longer Wars."

heat, thereby generating a paradigm change in what kinds of energy in human societies could be harnessed toward what ends, deep learning allows all the data accumulated by the proliferation of US sensors in recent decades to be repurposed toward substituting for human beings.[100] With deep learning, therefore, we should expect "more things in more places, more of the time."

**Massing Weapons**

First, a large extant discussion theorizes that AI may drive the rapid proliferation of lethal autonomous weapons systems (LAWS). On some views, the fielding of large autonomous drone swarms will fundamentally change warfare – as commonly envisioned, "swarm-on-swarm" warfare will involve the competitive attrition of many disposable units, in contrast to today's layered defense of high-value platforms like aircraft carriers.[101] While drone swarms have not yet appeared in modern militaries, deep learning solves a key bottleneck – that of a sufficient number of human pilots. In principle, after all, all the functions for which aircraft currently require a human being – integrating visual, audio, and signals information to execute the OODA loop ("Observe, Orient, Decide, Act"), maneuver to defeat air defenses, and make the decision to fire munitions – are all functions of a form which can be computed by deep learning's DNNs.

In terms of mass, with reference to competitions of pure attrition, partially or fully autonomous AI platforms would reduce a nation-state's need to recruit human bodies to fly planes or sweep cities. If even a highly finite pool of human aviators could iteratively pilot significant swarms of drones, this would shift the limiting factor in state ability to endure a protracted war to its stock of hardware, away from its stock of human bodies. As AI enables unmanned systems to increasingly represent the "tip of the spear," defense planners may come to count drones and their available domestic supply of AI chips the way they once counted soldiers and the available domestic recruiting pool of military-age adults.[102] Downstream, this could help demographically disadvantaged but technologically sophisticated nations like Japan compensate for low birth rates.[103]

---

[100] Joel Mokyr, ed., *The British Industrial Revolution: An Economic Perspective* (New York: Routledge, 2018), 20-1.

[101] Paul Scharre, "Robotics on the Battlefield Part II: The Coming Swarm," *CNAS*, October 2014, https://s3.amazonaws.com/files.cnas.org/documents/CNAS_TheComingSwarm_Scharre.pdf? (on cost-exchange ratios, see especially 20-3); Paul Scharre, "Robots at War and the Quality of Quantity," *War on the Rocks*, February 26, 2015, https://warontherocks.com/2015/02/robots-at-war-and-the-quality-of-quantity/; T. X. Hammes, "The Future of Warfare: Small, Many, Smart vs. Few & Exquisite?", *War on the Rocks*, July 16, 2014, https://warontherocks.com/2014/07/the-future-of-warfare-small-many-smart-vs-few-exquisite/; David Pinion, "The Navy and Marine Corps Need to Prepare for the Swarm of the Future," *War on the Rocks*, March 28, 2018, https://warontherocks.com/2018/03/the-navy-and-marine-corps-must-plan-for-the-swarm-of-the-future/; Joseph Hanacek, "The Perfect Can Wait: Good Solutions to the 'Drone Swarm' Problem," August 14, 2018, *War on the Rocks*, https://warontherocks.com/2018/08/the-perfect-can-wait-good-solutions-to-the-drone-swarm-problem/; Zachary Kallenborn and Philipp C. Bleek, "Drones of Mass Destruction: Drone Swarms and the Future of Nuclear, Chemical, and Biological Weapons," *War on the Rocks*, February 14, 2019, https://warontherocks.com/2019/02/drones-of-mass-destruction-drone-swarms-and-the-future-of-nuclear-chemical-and-biological-weapons/; Robert O. Work and Shawn Brimley, "20YY: Preparing for War in the Robotic Age," *CNAS*, January 2014, https://s3.amazonaws.com/files.cnas.org/documents/CNAS_20YY_WorkBrimley.pdf?.

[102] Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (W. W. Norton & Company, 2018).

[103] Todd Schneider, Gee Hee Hong, and Anh Van Le, "Land of the Rising Robots," *IMF*, June 2018, https://www.imf.org/external/pubs/ft/fandd/2018/06/japan-labor-force-artificial-intelligence-and-robots/schneider.pdf.

If autonomous drones enabled by deep learning are possible, states will face large incentives to use them. Owing to casualty aversion, the United States has sought to leverage advanced weapon systems to reduce the loss of its soldiers, and already makes extensive use of unmanned but human-piloted drones; the possibility of further reducing human deaths even in high-intensity air warfare would be a strong reason to seek adoption.[104] According to wargames research by Lin-Greenberg (2020), since the political costs of losing such drones would be low, widespread acquisition by states would likely lead to more frequent and longer in duration, but less escalatory wars.[105] Similarly, Gartzke (2019) also argues that remotely piloted vehicles will encourage frequent, drawn-out, low-intensity wars. After all, war typically proceeds until one side is either completely annihilated or discovers, before the other, after progressive absorption of warfare's costs, that its resolve is no longer sufficient to warrant willing continuation of the conflict. Insofar as wars increasingly risk only machines, however, they will only very slowly impose costs on involved nations, enabling glacially long comparisons of relative resolve on the battlefield, like two marathon runners using walking to carry out a contest of stamina.[106]

In addition, advanced LAWS, unfettered by human biological limits, would have several advantages over manned aircraft. Human cognitive ability varies between individuals, but within a fairly narrow band set by biology. In contrast, LAWS have no such ceiling, and will likely exhibit progressively more superior decision-making and reaction times.[107] Further, human biological limits mean aircraft cannot exceed survivable speeds as the G forces involved would kill the pilot. LAWS would suffer no such limit. In maturity, therefore, such systems would likely consistently defeat human pilots in dogfights.[108]

**Massing Propaganda**

Second, significant ink has been spilled over the use of deepfakes to generate propaganda at scale, or to feed an adversary false intelligence at a critical juncture, such as to manufacture an interstate crisis.[109] In

---

[104] Paul Scharre, "Robotics on the Battlefield, Part II: The Coming Swarm," CNAS, 2014, http://files.cnas.org.s3.amazonaws.com/documents/CNAS_TheComingSwarm_Scharre.pdf; ---, Army of None: Autonomous Weapons and the Future of War (New York: W. W. Norton & Company, 2018); Christian Brose, The Kill Chain: Defending America in the Future of High-Tech Warfare (New York: Hachette Books, 2020); Michael Horowitz, "Information-Age Weaponry and the Future Shape of Security in East Asia," GlobalAsia, 2011, https://globalasia.org/v6no2/cover/information-age-weaponry-and-the-future-shape-of-security-in-east-asia_michael-horowitz; Michael C. Horowitz, Joshua A. Schwartz, and Matthew Fuhrmann, "China Has Made Drone Warfare Global," Foreign Affairs, November 20, 2020, www.foreignaffairs.com/articles/china/2020-11-20/china-has-madedrone-warfare-global; Amy Zegart, "Cheap fights, credible threats: The future of armed drones and coercion," Journal of Security Studies 32.1 (2020): 6-46.

[105] Erik Lin-Greenberg, "Remote Controlled Restraint: The Effect of Remote Warfighting Technology on Crisis Escalation" (PhD diss., Columbia University, 2019); ---, "Wargame of Drones: Remotely Piloted Aircraft and Crisis Escalation" (unpublished manuscript), 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3288988.

[106] Erik Gartzke, "Blood and robots: How remotely piloted vehicles and related technologies affect the politics of violence," *Journal of Strategic Studies* (2019).

[107] Horowitz, "When speed kills," 769.

[108] Paul Scharre, "Robotics on the Battlefield, Part I: Range, Persistence, and Daring," *CNAS*, 2014.

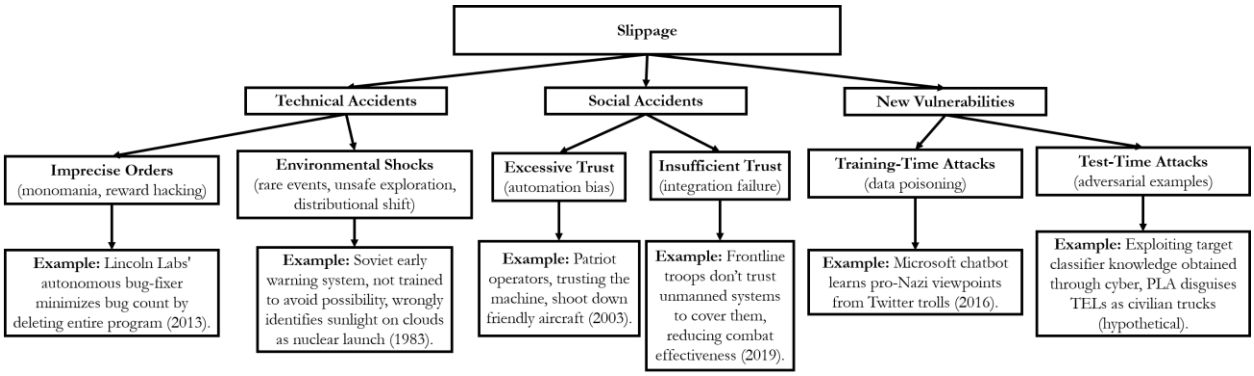
[109] Joe Littell, "Don't Believe Your Eyes (Or Ears): The Weaponization of Artificial Intelligence, Machine Learning, and Deepfakes," *War on the Rocks*, October 7, 2019, https://warontherocks.com/2019/10/dont-believe-your-eyes-or-ears-the-weaponization-of-artificial-intelligence-machine-learning-and-deepfakes/; for an interactive demonstration, see Donie O'Sullivan, "When seeing is no longer believing: Inside the Pentagon's race against deepfake videos," *CNN*, 2019, https://www.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/.

the future, current computational propaganda concerns may just be the tip of the iceberg, eclipsed by large-scale machine-generated of information that is targeted at particular individuals or subpopulations, evolved to maximally shape particular behaviors, and able to reach anyone as often as they have web access.[110]

Leveraging extensive progress in AI-empowered microtargeted advertising, governments could precisely manipulate human emotions, auto-generating text designed to appeal to the precise tastes, political leanings, and personality profile of a given citizen. OpenAI self-censored full publication of its GPT-2 language-generation model in 2019, for example, because it was concerned that generating close-to-human text would enable nefarious actors to proliferate disinformation; it is easy to imagine states pursuing the same for their own ends.[111] In the limit, an individual's entire information ecosystem – including ads for specific music, movies, or books, custom-generated news articles, and social media postings nominally associated with government organs, bloggers, or celebrities, could all be AI-generated. In such worlds, citizens of autocratic regimes would be encased inside informational cocoons of technological control, with the generated data ever increasing the AI system's ability to influence behavior and predict dissent.

### (c) Slippage

Besides accidents arising from deep learning's technical particulars *per se* ("misalignment"), integrating AI into military tasks also raises two additional categories of risk: social accidents related to interactions between AI and human beings, and new vulnerabilities related to adversaries attacking the new technology. I collectively refer to these accidents and vulnerabilities as "slippage," meaning any gap between intended and resulting effects of adopting the technologies. I diagram this below:




#### Social Accidents

*Excessive Trust*

Humans deployed with automated systems sometimes grow to assume the machine is always correct, worsening decision-making. Because the inner workings are somewhat incomprehensible, modern AI sometimes inspires a kind of reverence. As Go professional Fan Hui remarked after AlphaGo's move 37 against Lee Sedol, which led to decisive victory, "It's not a human move. I've never seen a human play this

[110] Matt Chessen, "The MADCOM Future," *The Atlantic Council*, 2017, https://www.atlanticcouncil.org/wp-content/uploads/2017/09/The_MADCOM_Future_RW_0926.pdf.

[111] Alec Radford, Jeffrey Wu, Dario Amodei, Daniela Amodei, Jack Clark, Miles Brundage, and Ilya Sutskever, "Better Language Models and Their Implications," *OpenAI*, February 14, 2019, https://openai.com/blog/betterlanguage-models/; Dipayan Ghosh and Ben Scott, "Digital Deceit: The Technologies Behind Precision Propaganda on the Internet," *New America*, January 23, 2018, https://www.newamerica.org/public-interest-technology/policypapers/digitaldeceit/, 26-8; Sarah Kreps and Miles McCain, "Not Your Father's Bots: AI Is Making Fake News Look Real," *Foreign Affairs*, https://www.foreignaffairs.com/articles/2019-08-02/not-your-fathers-bots.

move. So beautiful."[112] Broad empirical evidence exists for this phenomenon.[113] Known as "automation bias," this was the cause of several fratricides by Patriot batteries in 2003, during the invasion of Iraq – operators trusted the targeting solutions generated by the batteries, resulting in firing on friendly aircraft.[114]

*Insufficient Trust*

Conversely, humans with modern AI systems, not understanding them, sometimes also exhibit the opposite reaction, declining to trust these systems altogether. For example, Macdonald and Schneider (2019) find strong survey evidence for preferences for manned over unmanned aircraft among American ground fires controllers.[115] Lack of trust can lead to units in the field being combat ineffective, as well as services in general failing to genuinely integrate AI into doctrinal and operational practice.[116]

**New Vulnerabilities**

In military affairs, adopting new capabilities means adopting new vulnerabilities, as the addition of any technology logically expands the attack surface available to adversaries. For example, the information revolution also spawned the need for cybersecurity.[117] We can divide AI vulnerabilities into two types: training-time attacks and test-time attacks.

---

[112] Cade Metz, "The sadness and beauty of watching Google's AI play Go," *Wired*, March 11, 2016, https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/.

[113] Kate Goddard, Abdul Roudsari, and Jeremy C. Wyatt, "Automation bias: a systematic review of frequency, effect mediators, and mitigators," *Journal of the American Medical Informatics Association* 19 (2012): 121-7; Linda J. Skitka, Kathleen L. Mosier, and Mark Burdick, "Does automation bias decision-making?", *International Journal of Human-Computer Studies* 51 (1999): 991-1006.

[114] John K. Hawley, "Patriot Wars: Automation and the Patriot Air and Missile Defense System," *CNAS*, January 2017, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Report-EthicalAutonomy5-PatriotWars-FINAL.pdf?; Scharre, *Army of None*, 138-179. See also Wiener, "Some Moral and Technical Consequences of Automation," 1357; and Geist and Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?", 18.

[115] Julia Macdonald and Jacquelyn Schneider, "Battlefield Responses to New Technologies: Views from the Ground on Unmanned Aircraft," *Security Studies* 28.2 (2019), 216-49.

[116] Margarita Konaev, Tina Huang, and Husanjot Chahal, "Trusted Partners: Human-Machine Teaming and the Future of Military AI," CSET, https://cset.georgetown.edu/research/trusted-partners/. Illustratively, history contains many cases of new technology hobbling, rather than helping, militaries unable to understand and trust it. For example, during the Chinese Civil War, the United States supplied advanced weapon systems to its Nationalist allies in hopes of assisting defeat of the Communists. This backfired, however, as the Nationalist military lacked the logistical, doctrinal, and organizational ability to make proper use of advanced American weaponry, leading to sharply reduced military effectiveness relative to other Nationalist units not supplied with higher-end American weapons at all. For example, many Communist forces had never before even seen a tank, but Nationalist units made their American-supplied tanks immobile by choosing to dig a trench around them, then drove them over their own soldiers in a disorganized panic when attacked. An elite Nationalist division drove its tanks during heavy rain into a swamp, became stuck in the mud, and was completely eliminated by Communist forces. This problem was so severe that Chiang Kai-shek eventually banned his forces from using advanced weapons entirely in 1947. Victor Cheng, "Modern War on an Ancient Battlefield: The Diffusion of American Military Technology and Ideas in the Chinese Civil War, 1946-1949," *Modern China* 35.1 (2009), 38-64. For an excellent fictional depiction of this dynamic, see Arthur C. Clarke, "Superiority," *The Magazine of Fantasy and Science Fiction* 2.4 (1951), 3-12. Available online: http://nob.cs.ucdavis.edu/classes/ecs153-2019-04/readings/superiority.pdf. To quote the main character, "We were defeated by one thing only - by the *inferior* science of our enemies."

[117] Jacquelyn Schneider, "The capability/vulnerability paradox and military revolutions: Implications for computing, cyber, and the onset of war," *Journal of Strategic Studies* 42.6 (2019): 841-63.

As the name suggests, *training-time attacks* occur during the model training phase, when the deep neural net is in the process of learning how to, say, separate cats from dogs. Here, most often discussed is "data poisoning," where attackers add to training data to alter later model behavior. These attacks can be fairly sophisticated – for example, researchers have found adding just a single image to a training set can allow the attacker to control the identity of some target person, in the view of the classifier.[118]

*Test-time attacks* occur after training, when the AI is attempting to use its model to classify data. Here, most research focuses on adversarial examples, which are input data with human-imperceptible modifications intended to induce misclassification. This attack is possible because a deep neural network's input-output mappings are significantly discontinuous – that is, in the n-dimensioned space of possible images, the learned model does not consider inputs close to those it would classify as "cat" to also be cats. Instead, moving a human-imperceptible amount in that space might lead the model, for example, to believe with high confidence that the input was a tortoise instead.[119] In contrast, even with significantly larger perturbations than those necessary to fool deep learning models, humans would consider most visual inputs close to those intuitively classified as cats to be, at the least, strongly catlike.

Importantly, attacks utilizing adversarial examples do not require digital access to the model, as long as the targeted AI system intakes perturbable physical data. For example, an attacker could imperceptibly modify a pop song to contain audio data recognized by the AI system as voice commands, or apply subtle makeup to their face to cause the AI system to believe they are someone with access rights.[120] Further, even in "black box" scenarios where the attacker has no prior knowledge about the model, the attacker may still be able to generate adversarial examples either by training against known models they believe to be similar (e.g., an outdated, but now public iteration of the same classifier).[121]

Considerable research effort explores how to defend machine learning models against adversarial examples, such as by oneself generating and then directly training against such examples ("adversarial training"), or by smoothing the model's input-output mappings through outputting probabilities, rather than hard

---

[118] Ali Shafahi, W. Ronny Huang, Mahyar Najibi, Octavian Suciu, Christoph Studer, Tudor Dumitras, and Tom Goldstein, "Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks," NeurIPS, 2018, https://papers.nips.cc/paper/7849-poison-frogs-targeted-clean-label-poisoning-attacks-on-neural-networks.pdf. Chen et al. previously found attackers can also create difficult-to-detect backdoors into deep learning-based authentication systems, even without knowledge of the model or training set, if permitted to inject around 50 poisoning samples. Xinyun Chen, Chang Liu, Bo Li, Kimberly Lu, and Dawn Song, "Targeted Backdoor Attacks on Deep Learning Systems Using Data Poisoning," arXiv, December 15, 2017, https://arxiv.org/pdf/1712.05526.pdf. See discussion of defenses against data poisoning attacks at Jacob Steinhardt, Pang Wei Koh, and Percy Liang, "Certified Defenses for Data Poisoning Attacks," NIPS, 2017, https://papers.nips.cc/paper/6943-certified-defenses-for-data-poisoning-attacks.pdf.

[119] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, Rob Fergus, "Intriguing properties of neural networks," arXiv, February 19, 2014, https://arxiv.org/pdf/1312.6199.pdf.

[120] Alexey Kurakin, Ian J. Goodfellow, and Samy Bengio, "Adversarial Examples in the Physical World," arXiv, February 11, 2017, https://arxiv.org/pdf/1607.02533.pdf.

[121] Alexey Kurakin, Ian Goodfellow, Samy Bengio, Yinpeng Dong, Fangzhou Liao, Ming Liang, Tianyu Pang, Jun Zhu, Xiaolin Hu, Cihang Xie, Jianyu Wang, Zhishuai Zhang, Zhou Ren, Alan Yuille, Sangxia Huang, Yao Zhao, Yuzhe Zhao, Zhonglin Han, Junjiajia Long, Yerkebulan Berdibekov, Takuya Akiba, Seiya Tokui, and Motoki Abe, "Adversarial Attacks and Defences Competition," *arXiv*, March 31, 2018, https://arxiv.org/pdf/1804.00097.pdf; Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z. Berkay Celik, and Ananthram Swami, "Practical Black-Box Attacks against Machine Learning," *arXiv*, March 19, 2017, https://arxiv.org/pdf/1602.02697.pdf/.

decisions ("defensive distillation"). Such defenses can still be defeated if the attacker holds an advantage in computational resources, however.[122]

### (3) Strategic Effects

Finally, how might these tactical effects impact the US-China balance of power? The conventional wisdom argues that AI may be a critical technology in US-China competition, assessing that the boost from harnessing AI might suffice for China to leapfrog US power. For former Deputy Secretary of Defense Robert Work and Greg Grant, just as the United States "offset" the Soviet Union's conventional overmatch in Europe first with tactical nuclear weapons, then later with long-range precision strike, so too might China offset American advantages by winning in AI. In their view, "any objective assessment must at least consider the possibility that the U.S. Joint Force is close to becoming the victim of a deliberate, patient, and robustly resourced military-technical offset strategy. ... Chinese military thinkers believe AI likely will be the key to surpassing the U.S. military as the world's most capable armed force."[123]

Similarly, according to Elsa Kania, Chinese elites indeed see AI as a "leapfrog" technology that could help China surge past the United States; Greg Allen concurs.[124] Fu Ying, who chairs both the PRC National People's Congress Foreign Affairs Committee and Tsinghua University's Center for International Strategy and Security, AI may "set[] off a new round of the rise and fall of the great powers."[125] Similarly, Graham Allison assesses that China may "beat America to AI supremacy," and "is currently on a trajectory to overtake the United States in the decade ahead."[126] James Johnson argues AI will "redefine and transform the status quo in military-use technology," promoting a "shift to Sino-American bipolarity."[127]

AI's effects on the US-China balance of power have general theoretical significance. Many analysts argue emerging technologies represent opportunities for weaker powers to leapfrog US strength; others are equally skeptical.[128] Security studies has long maintained an ambivalent relationship with technology as a

---

[122] Ian Goodfellow, Nicolas Papernot, Sandy Huang, Rocky Duan, PIeter Abbeel, and Jack Clark, "Attacking Machine Learning with Adversarial Examples," *OpenAI*, February 24, 2017.

[123] Robert O. Work and Greg Grant, "Beating the Americans at Their Own Game: An Offset Strategy with Chinese Characteristics," *CNAS*, June 6, 2019, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Report-Work-Offset-final-B.pdf?, 14.

[124] Elsa B. Kania, "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power," *CNAS*, November 28, 2017, https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power; Gregory C. Allen, "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security," *CNAS*, February 6, 2019, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

[125] Fu Ying, trans. Brian Tse and Jeffrey Ding, "A Preliminary Analysis of the Impact of AI on International Relations," *Quarterly Journal of International Politics* 4 (2019): 1-18. Translation available at https://chinai.substack.com/p/chinai-67-fu-ying-on-ai-the-international.

[126] Graham Allison, "Is China Beating America to AI Supremacy?," *The National Interest*, December 22, 2019, https://nationalinterest.org/feature/china-beating-america-ai-supremacy-106861.

[127] James Johnson, "The end of military-techno Pax Americana? Washington's strategic responses to Chinese AI-enabled military technology," *The Pacific Review* (2019), https://www.tandfonline.com/action/showCitFormats?doi=10.1080/09512748.2019.1676299, 7, 11.

[128] Todd S. Sechser, Neil Narang, and Caitlin Talmadge, "Emerging technologies and strategic stability in peacetime, crisis, and war," *Journal of Strategic Studies* 42.6 (2019): 727-35. This idea dates back to Modelski.

causal variable – is technology almost epiphenomenal, the product of conscious investment driven by the typical forces of international relations, or is it capable of being an independent force, with quasi-random discoveries carrying the significance of exogenous shocks? Arguably, AI is a useful test of these theories.

In my view, however, the case for Chinese advantage thus far lacks conceptual precision. First, due in part to definitional confusions, what actual inputs are required for modern states to acquire AI have not been precisely discussed. Second, even after states acquire AI, there is no universal conversion rate from "more AI" to "more power." As having "more AI" is unlike having "more tanks," evaluating who AI advantages requires theory beyond bean-counting. Without further conceptual glue, assessments that China possesses a "data advantage" or that American talent still dominates remain only standalone facts of uncertain effect and magnitude. Thus, correspondingly, I divide analysis into two parts, looking at differences in:

> ● **Inputs.** Put colloquially, **"**how much AI" can each state get? Deep learning's key precursors are task-specific data, computational power, and AI talent. Can either state simply buy as much of those precursors as they like on the open market, in which case access to inputs may be effectively symmetric, or are there further complications?
>
> ● **Leverage.** Once acquired, "how useful is that AI" to each state in terms of the relative balance? Critically, where, when, and how much does access to AI's tactical effects, for either state, translate into real strategic effects? Notably, these effects need not be symmetric, as the United States and China have different strategic circumstances and needs.

### (a) Inputs

"How much AI" can the US and China get? New technologies generate new inputs to power: for example, the advent of railroads, internal combustion engines, and nuclear weapons dramatically increased the value of steel, oil, and uranium, respectively.[129] A state's endowment of these resources or their precursors, previously irrelevant, became variously important for its security. So, too, did that state's policies, or lack thereof, for acquiring and exploiting those inputs.

In the case of AI, recent deep learning advances have been driven by three key factors: algorithmic advances by leading researchers, hardware improvements via inexpensive parallel-processing GPUs followed by advanced AI chips, and expanding data availability in an exploding number of fields.[130] Consequently, we can ask about the balance of access to talent, hardware, and data in turn.

#### Talent

The United States likely possesses a mild to moderate talent advantage. First, consider pure manpower. While the PRC's AI emphasis led to dozens of universities setting up AI degree programs, much of the newly educated talent has left China. Of the 2,800 Chinese accepted to prominent machine learning conference NeurIPS over the last decade, for example, three-quarters are working outside China. Further, 85% of those departees headed to the United States. [131] According to recent data on 1,999 US AI PhD

---

[129] On railroads and steel, see Goldman and Andres, "Systemic effects of military innovation and diffusion," 116. On the internal combustion engine and oil, see W. G. Jensen, "The Importance of Energy in the First and Second World Wars," *The Historical Journal* 11 (1968): 538-54. On uranium, see R. Scott Kemp, "The Nonproliferation Emperor Has No Clothes: The Gas Centrifuge, Supply-Side Controls, and the Future of Nuclear Proliferation," *International Security* 38 (2014): 39-78, especially 41-4.

[130] Goodfellow et al., *Deep Learning*, 18-26.

[131] Joy Dantong Ma, "China's AI Talent Base Is Growing, and then Leaving," *MacroPolo*, July 30, 2019, https://macropolo.org/chinas-ai-talent-base-is-growing-and-then-leaving/.

graduates, over 80 percent of internationally originating US-trained AI PhDs stay in the United States after graduation, while the "vast majority" of Chinese-trained talent has left China. In fact, over 90 percent of Chinese students in this data choose to stay in the United States.[132]

Similarly, Tencent has reported that China contains around 39,000 AI researchers, half of the United States' 78,000 researchers. Notably, the total talent pool is quite small – Tencent notes an upper bound at around 300,000 qualified AI researchers, but this figure also includes non-expert members of technical AI teams. According to an analysis of LinkedIn data by J. F. Gagne, the total pool of PhD-educated AI researchers numbers only 22,000, with China able to draw only one-fourth of new researchers trained abroad compared to the United States.[133] Most US data scientists have over a decade of experience, while 40 percent of those in China have less than half a decade of the same.[134]

Second, consider research quality. According to Stanford's 2019 AI Index, US publications in AI retain 50% greater influence than those of China by Field-Weighted Citation Impact (FWCI), a regionally-adjusted measure of citation frequency.[135] Similarly, according to McKinsey, while China has become increasingly prolific in publishing a large raw number of papers, if adjusted by H-index, a measure of influence, China (168) lags even the UK (190) and has less than half the influence score of the United States United States (373).[136] Subjectively, US companies have dominated headlining AI advances, such as with AlphaGo defeating Lee Sedol and OpenAI's development of NLP model GPT-2. The dominant platforms used in AI research, TensorFlow and PyTorch, are also American.[137]

Further, talent concentrations are likely self-sustaining. Open Economy Politics (OEP), the dominant research paradigm in international political economy (IPE), has long adopted the assumption of constant returns to scale, allowing analysis of trade across industries. Many goods, however, exhibit high intra-industry trade, indicating increasing returns to production. If AI is such a good, the key economic implication is that despite the wide availability of certain factors of production such as software, the

---

[132] Remco Zwetsloot, James Dunham, Zachary Arnold, and Tina Huang, "Keeping Top AI Talent in the United States: Finding and Policy Options for International Graduate Student," CSET, December 2019, https://cset.georgetown.edu/wp-content/uploads/Keeping-Top-AI-Talent-in-the-United-States.pdf. Of course, the future is unwritten. Notably, Zwetsloot et al. implicate recently increased barriers to immigration as potentially threatening these trends.

[133] J. F. Gagne, Fedor Karmanov, and Simon Hudson, "Global AI Talent Report 2018," *jfgagne*, https://jfgagne.ai/talent/; J. F. Gagne, Grace Kiser, and Yoan Mantha, "Global AI Talent Report 2019," jfgagne, https://jfgagne.ai/talent-2019/; Joy Dantong Ma, "The AI Race Is Wide Open, If America Remains Open," *Macro Polo*, April 15, 2019, https://macropolo.org/us-china-ai-race-talent/.

[134] Dominic Barton, Jonathan Woetzel, Jeongmin Seong, and Qinzheng Tian, "Artificial Intelligence: Implications for China," *McKinsey*, April 2017, https://www.mckinsey.com/featured-insights/china/artificial-intelligence-implications-for-china, 5-8.

[135] Raymond Perrault, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles, "The AI Index 2019 Annual Report," Human-Centered AI Institute, December 2019, https://hai.stanford.edu/sites/g/files/sbiybj10986/f/ai_index_2019_report.pdf, 18.
[136] Barton, et al., "Artificial Intelligence," 5-8.

[137] Imbrie et al., "The Question of Comparative Advantage in Artificial Intelligence," 11.

distributional benefits of AI may nonetheless not widely diffuse, but rather concentrate in a small number of countries.[138]

Indeed, Avi Goldfarb and Daniel Trefler argue that AI contains several economies of scale. Data manifests economies of scale by enabling better product, which in turn attracts more customers, who provide more data. Building AI capabilities requires large fixed costs: while much research is open-source, salaries for high-quality AI researchers often reach into the millions.[139] Finally, AI investments exhibit "economies of scope," where building multiple products is cheaper than building each separately. For example, the Alphabet subsidiaries Google, YouTube, Android, and Waymo all benefit from their pooled in-house AI capabilities, and data, talent, and software also passes easily between these firms.[140]

Why do these benefits require geographic colocation? The central insight of "new economic geography," as first conceived of by Paul Krugman in the 1990s, is very roughly that economic activity tends to have good reasons for geographic aggregation (e.g., cities exist), and that these reasons then weigh against other factors encouraging diffusion. Despite intense competition and the need to pay high wages and property costs, for example, firms continue to stubbornly locate in Silicon Valley and New York, along with in other persistent regional clusters globally.[141]

For AI, empirically, even today, capabilities have become concentrated in a small handful of companies in a small handful of locations. Two reasons arise: the costs of transmitting data across national boundaries, as regulatory rules in different jurisdictions arise, and perhaps more importantly the local diffusion of tacit engineering knowledge, transmissible primarily through face-to-face engineer contact over the course of an organization's regular functioning.[142] In fact, talent location is evidently so important that companies have empirically been willing to set up new branches near key hires, so that professors can retain their academic jobs will working for some new division. Examples here include Raquel Urtasun, both of the University of Toronto and Uber, Russ Salakhutdinov of Carnegie Mellon University and Apple, and Yann LeCun, of NYU and Facebook. Google's DeepMind opened in London then expanded to Edmonton exactly because Demis Hassabis and Richard Sutton lived in those locations, respectively. Each was afforded this exorbitant privilege by their respective company.[143]

---

[138] David A. Lake, "Open economy politics: A critical review," *The Review of International Organizations* 4 (2009), 236-7.

[139] Avi Goldfarb and Daniel Trefler, "AI and International Trade," forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 8-9. Available online: https://www.nber.org/books/agra-1.

[140] Ibid., 9-10.

[141] The literature is large. For retrospectives, see Paul Krugman, "The New Economic Geography, Now Middle-aged," *Regional Studies* 45 (2010): 1-7; and Masahisa Fujita and Paul Krugman, "The new economic geography: Past, present and the future," *Papers in Regional Science* 83 (2005): 139-64. For the foundational work, see Paul Krugman, *Geography and Trade* (MA: MIT Press, 1991); and Masahisa Fujita, Paul Krugman, and Anthony J. Venables, *The Spatial Economy: Cities, Regions, and International Trade* (MA: MIT Press, 2001).

[142] Ibid., 29, 10-11; Catherine Durnell Cramton, "Insights for Culture and Psychology from the Study of Distributed Work Teams," in *Handbook of Advances in Culture and Psychology, Volume 6*, ed. Michele J. Gelfand, Chi-yue Chiu, and Ying-yi Hong (UK: Oxford University Press, 2015). Gilli and Gilli also note this in their study of drone-related software. See Gilli and Gilli, "The Diffusion of Drone Warfare?", 77.

[143] Goldfarb and Trefler, "AI and International Trade," 10-12.

This talent deficit will likely impact China's relative ability to leverage AI for military ends, as even with readily available compute and data, model architecture depends strongly on intelligent human input, and AI research on military applications specifically is unlikely to be openly shared.[144] This is especially so because a key difficulty in training AI for military tasks is that the chaos of the battlefield is long-tailed – that is, much of the distribution consists of individually rare events. Predictably, classifiers relying on large datasets struggle with long-tailed phenomena, and require careful tuning by human engineers.[145]

**Hardware**

Second, how does Chinese hardware compare to that of the United States? While much software is open-source and Chinese skill at IP theft is legendary, one source of evidence is to observe what the actual state of hardware development is on both sides of the Pacific. Notably, China is significantly behind in two key inputs into AI development: semiconductors and data centers.[146]

First, while China has made progress in producing memory chips, as well as analog and compound semiconductors, it has had much more difficulty specifically with semiconductor foundries.[147] Within semiconductor manufacturing equipment (SME), two bottlenecks most constrict China – photolithography equipment, the manufacture of which is dominated by the Netherlands and Japan, and deposition, etching, and process control equipment, the manufacture of which is dominated by US and Japanese companies. Overall, over 90 percent of SME firms are in the United States, the Netherlands, and Japan.[148] Since nearly all high-end AI chips are also made in Taiwan, the United States, and South Korea, the United States, with allied cooperation, is consequently able to choke off Chinese hardware supplies for AI. For example, in October 2018, US sanctions on Chinese chip company Fujian Jinhua forced a halt in output the following March.[149]

Simultaneously, Chinese efforts to create indigenous capacity have not succeeded. China's Semiconductor Manufacturing International Corporation (SMIC) has consistently lagged the Taiwan Semiconductor Manufacturing Company (TSMC), the world's current largest semiconductor foundry. SMIC had plans to switch to a 14nm process for production in 2019, which TSMC has been using at scale since 2015. Since TSMC is the market leader, it has been able to command high prices for its services, while SMIC has

---

[144] Joel Hestness, Sharan Narang, Newsha Ardalani, Gregory Diamos, Heewoo Jun, Hassan Kianinejad, Md. Mostofa Ali Patwary, Yang Yang, and Yanqi Zhou, "Deep Learning Scaling is Predictable, Empirically," December 1, 2017, *arXiv*, https://arxiv.org/pdf/1712.00409.pdf, 1.

[145] Grant Van Horn and Pietro Perona, "The Devil is in the Tails: Fine-grained Classification in the Wild," *arXiv*, September 5, 2017, https://arxiv.org/pdf/1709.01450.pdf; Yu-Xiong Wang, Deva Ramanan, and Marital Hebert, "Learning to Model the Tail" (paper presented at the 31st Conference on Neural Information Processing Systems, Long Beach, California, 2017). Available online: https://papers.nips.cc/paper/7278-learning-to-model-the-tail.pdf.

[146] I discuss US hardware advantage at much greater length in the third essay.

[147] Dan Wang, "Catching Up In Chips," *Gavekal*, October 1, 2018.

[148] Saif M. Khan, "Maintaining the AI Chip Competitive Advantage of the United States and its Allies," *CSET*, December 2019, http://cset.georgetown.edu/wp-content/uploads/CSET-Maintaining-the-AI-Chip-Competitive-Advantage-of-the-United-States-and-its-Allies-20191206.pdf.

[149] Andrew Imbrie, Elsa B. Kania, and Lorand Laskai, "The Question of Comparative Advantage in Artificial Intelligence: Enduring Strengths and Emerging Challenges for the United States," *CSET*, January 20, http://cset.georgetown.edu/wp-content/uploads/CSET-The-Question-of-Comparative-Advantage-in-Artificial-Intelligence-1.pdf, 6-7; Kathrin Hille, "Trade war forces Chinese chipmaker Fujian Jinhua to halt output," *Financial Times*, January 28, 2019, https://www.ft.com/content/87b5580c-22bf-11e9-8ce6-5db4543da632.

required infusions of government subsidies.[150] Breaking into technically complex, consolidated industries is naturally challenging, and China lacks indigenous talent. Empirically, Chinese efforts here have consistently leaned on state-owned enterprises (SOEs), with such unconvincing results that one industry expert has remarked "I think what they are doing in fabrication is another Great Leap Forward."[151] Despite disproportionate government equity injections, Chinese semiconductor firm profitability remains behind European, Japanese, Korean, Taiwanese, and American peers.[152] Consequently, Saif Khan estimates China will "very likely" fail to create an indigenous photolithography industry inside a decade.

Further, US export controls explicitly seek to keep Chinese semiconductor manufacturing capabilities two generations behind the state of the art, requiring China to make up the gap natively.[153] Catching up will likely a difficult task for China – semiconductor foundries form a natural oligopoly as a corollary to Moore's Law, as the parallel continuous increase in foundry costs causes profitability to accrue only to leading companies. The entry barrier is thus very high – according to Rock's Law (sometimes called Moore's Second Law), the cost of a semiconductor fabrication plant should double every four years.[154]

Second, one further way to operationalize hardware is to compare data centers. Since 2012, the most compute-intensive AI training runs have doubled in compute usage, astonishingly, every 3.5 months.[155] As such, continued AI progress will likely require increasingly massive investments in compute, frequently actualized as "hyperscale" data centers which exploit efficiencies of scale to power machine learning's increasingly greedy demands.[156] This is especially so as tasks with greater subjective difficulty have, empirically, been increasingly amenable to parallelization. Within supervised learning, for example, training on ImageNet's 14 million images of all sorts has been more parallelizable than with MNIST, a relatively simple database of handwritten letters. Similarly, within reinforcement learning, Dota 5v5 has been more parallelizable than Atari Pong.[157] This will likely increase both the monetary cost and hardware

[150] Wang, "Catching Up In Chips"; Tim Hwang, "Computational Power and the Social Impact of Artificial Intelligence," *arXiv*, March 2018, https://arxiv.org/abs/1803.08971.

[151] Douglas B. Fuller, "Growth, Upgrading, and Limited Catch-Up in China's Semiconductor Industry," in *Policy, Regulation and innovation in China's Electricity and Telecom Industries*, eds. Loren Brandt and Thomas G. Rawski (Cambridge University Press, 2019), 297.

[152] Andrea Andrenelli, Julien Gourdon, Yuki Matsumoto, Taku Nemoto, Jehan Sauvage, and Christian Steidl, "Measuring distortions in international markets: The semiconductor value chain," *OECD Trade Policy Papers* 234 (2019), https://www.oecd-ilibrary.org/docserver/8fe4491d-en.pdf?, 73.

[153] "Rapid Advances in China's Semiconductor Industry Underscore Need for Fundamental U.S. Policy Review," *GAO*, April 2002, https://www.gao.gov/new.items/d02620.pdf.

[154] Karl Rupp and Siegfried Selberherr, "The Economic Limit to Moore's Law," *Proceedings of the IEEE* 98 (2010): 351-3. Available online: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5415663.

[155] Dario Amodei and Danny Hernandez, "AI and Compute," *OpenAI*, May 16, 2018, https://openai.com/blog/ai-and-compute/.

[156] Scott Fulton, "How hyperscale data centers are reshaping all of IT," *ZDNet*, April 5, 2019, https://www.zdnet.com/article/how-hyperscale-data-centers-are-reshaping-all-of-it/. In fact, OpenAI moved away in early 2019 from being purely nonprofit expressly to raise more capital for investment in compute, which it saw as necessary to continuing research progress in AI. See "OpenAI LP," *OpenAI*, March 11, 2019, https://openai.com/blog/openai-lp/.

[157] Sam McCandlish, Jared Kaplan, and Dario Amodei, "How AI Training Scales," OpenAI, December 14, 2018, https://openai.com/blog/science-of-ai/.

reliance of future research – while the largest runs in 2018 relied on hardware costing millions of dollars, continued growth at the same rate implies moving from AlphaGo Zero's $10 million to perhaps $5-6 billion dollar experiment sizes within a half decade.[158] Notably, OpenAI's training run of GPT-3 was so expensive that although a minor bug was discovered after starting, the company could not restart the process and had to accept a slight inefficiency.[159]

By the data center metric, the United States is well ahead. As of December 2018, the United States was home to 40% of the world's hyperscale data centers, well ahead of China's 8% share. Amazon and Google together opened almost half of new data centers created in 2018.[160] Quantity aside, data centers also operate at differing levels of efficiency, measured by Power Usage Effectiveness (PUE), the ratio of total energy used by the data center to energy successfully used for computing. As such, a PUE of 1 would represent perfect efficiency, with zero overhead. Most Chinese data centers operate at between 2.2 and 3.0 PUE, consuming 2 percent of total Chinese electricity consumption, despite government efforts to encourage greater efficiencies, including a ban on data centers with over 1.5 PUE from operating in Beijing. In contrast, the average US data center operates at 1.82 PUE.[161] As of May 2019, the average Google data center operated at 1.11 PUE, with the most efficient as low as 1.07.[162] Baidu declined to provide a 2017 Greenpeace study with PUE data.[163]

**Data**

Finally, let us turn to data. One common metaphor is that data is the new oil.[164] The comment that typically follows is that China has more data, and that AI therefore advantages China, which possesses both a larger population size and relatively lax privacy laws allowing the government to access the data that population

---

[158] Amodei and Hernandez, "AI and Compute"; Ryan Carey, "Interpreting AI Compute Trends," *AI Impacts*, July 10, 2018, https://aiimpacts.org/interpreting-ai-compute-trends/.

[159] Brown et al., "Language Models are Few-Shot Learners."

[160] "Hyperscale Data Center Count Jumps to 430; Another 132 in the Pipeline," *Synergy Research Group*, January 9, 2019, https://www.srgresearch.com/articles/hyperscale-data-center-count-jumps-430-mark-us-still-accounts-40.

[161] Ying Zhao, "How LEED helps the growth of data centers in China," *GBCI*, May 24, 2018, http://www.gbci.org/how-leed-helps-growth-data-centers-china; Coco Liu, "For China's Massive Data Centers, a Push to Cut Energy and Water Use," *Environment 360*, August 22, 2016, https://e360.yale.edu/features/for-chinas-massive-data-centers-a-push-to-cut-energy-and-water-use.

[162] "Efficiency: How we do it," *Google*, accessed May 7, 2019, https://www.google.com/about/datacenters/efficiency/internal/.

[163] "Clicking Clean: Who is Winning the Race to Build a Green Internet?", *Greenpeace*, 2017, https://secured-static.greenpeace.org/austria/Global/austria/dokumente/Clicking%20Clean%202017.pdf; Yevgeniy Sverdlik, "Pollution in China Makes Free Cooling Difficult for Baidu," *DataCenter Knowledge*, August 3, 2015, https://www.datacenterknowledge.com/archives/2015/08/03/china-data-center-operators-struggle-with-pollution.

[164] For example, see Joris Toonders, "Data Is the New Oil of the Digital Economy," *WIRED*, 2014, https://www.wired.com/insights/2014/07/data-new-oil-digital-economy/; "The world's most valuable resource is no longer oil, but data," *The Economist*, May 6, 2017, https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data; though see also disagreement at Jocelyn Goldfein and Ivy Nguyen, "Data is not the new oil," *TechCrunch*, March 27, 2018, https://techcrunch.com/2018/03/27/data-is-not-the-new-oil/; and Antonio Garcia Martinez, "No, Data is Not the New Oil," *WIRED*, February 26, 2019, https://www.wired.com/story/no-data-is-not-the-new-oil/.

generates.[165] In the most well-known advancement of this narrative, Kai-Fu Lee has argued that AI is "turning China into the Saudi Arabia of data: a country that suddenly finds itself sitting atop stockpiles of the key resource that powers this technological era."[166] Similarly, in an influential report for the Center for a New American Security, Elsa Kania cites China's "massive amounts of data" as potentially enabling a "rapid rise" for China in AI, citing a China Electronic Information Industry Development (CCID) study showing China will have "20% of the world's data by 2020 and 30% by 2030."[167] Indeed, the International Data Corporation projects that China's share of global data will grow from 23.4% in 2018 to 27.8% in 2025, outpacing average global growth by 3%.[168] Simultaneously, they project the US share will decline from 21% to 17.5%.[169]

Advocates of this view are not unaware that other factors go into AI; rather, they simply believe that AI is the most important driver. Baidu CEO Qi Lu, for example, has argued that data is paramount due to being the "primary means of production."[170] Similarly, for Lee, AI has now entered an the "age of implementation," with the most economic purchase to be gained via applying existing AI insights, rather than discovering new ones. As such, AI today is comparable to the task of mass electrification – companies transforming themselves following the discovery of electricity did not need to achieve new breakthroughs, but merely competently implemented Edison's discovery.[171] In his pithy formulation, "In deep learning, there's no data like more data … Given much more data, an algorithm designed by a handful of mid-level AI engineers usually outperforms one designed by a world-class deep-learning researcher."[172]

This view is not without merit – after all, large dataset availability has been a key ingredient in powering the recent resurgence of artificial intelligence.[173] However, a number of factors significantly mitigate this

[165] Barton et al., "Artificial Intelligence," 5.

[166] Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt, 2018), 55.

[167] Elsa B. Kania, "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power," November 2017, *CNAS*, https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power, 11, 33.

[168] David Reinsel, Lianfeng Wu, John F. Gantz, and John Rydning, "The China Datasphere: Primed to Be the Largest Datasphere by 2025," *IDC*, January 2019, https://www.seagate.com/files/www-content/our-story/trends/files/data-age-china-idc.pdf.

[169] John F. Gantz, David Reinsel, and John Rydning, "The U.S. Datasphere: Consumers Flocking to Cloud," *IDC*, January 2019, https://www.seagate.com/files/www-content/our-story/trends/files/data-age-us-idc.pdf. (Graphic to be adapted.) David Reinsel, John Gantz, and John Rydning, "The Digitization of the World From Edge to Core," *Seagate*, 2018, https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf. Here, EMEA is Europe, the Middle East, and Africa, and APJxC is the Asia-Pacific excluding China.

[170] As cited in Ding, "Deciphering China's AI Dream," 28.

[171] Lee, *AI Superpowers*, 86.

[172] Ibid., 14.

[173] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM* 60 (2017): 84-90. Available online: https://www.cs.toronto.edu/~kriz/imagenet_classification_with_deep_convolutional.pdf.

advantage.[174] First, data generalizes poorly. For example, models trained in Wikipedia's standard English have not been able to understand Twitter's casual slang.[175] China's population size likely makes it the world-leading authority on performing, say, facial recognition of Chinese faces, but this does not enable its algorithms to better scan battlefields for tanks or identify cyber-vulnerabilities in American systems.[176] As such, China's data advantage is narrow and uneven – Chinese AI-powered mobile apps might certainly have more data about how to get more play time out of consumers, for example, but proximate benefits are likely economic, not military. Data is not oil, but rather kerosene, diesel, jet fuel, and so on.[177]

In fact, for military-relevant applications, it seems plausible that the United States is the party with the data advantage.[178] During the Cold War, the United States invested extensively in ASW capabilities to locate and track Soviet SSBNs, resulting in sequential breakthroughs in passive sonar and signals processing. These included an underwater network of hydrophones in the '50s, followed by attack submarines with both bow mounted and towed arrays in the '60s and '70s, which forced the Soviets to withdraw their SSBNs defensively into near-shore bastions. To track land-based mobile assets, the United States developed a constellation of satellites possessing synthetic aperture radar, which has become capable over the last two decades of detecting moving targets. Stealthy penetrating UAVs and unattended ground sensors further provide relevant military data about enemy asset movements.[179] By raw numbers, the United States operates 901 satellites to China's 299; while China has been significantly investing in augmenting its capabilities, the US still has a significant overall lead.[180] The geography of the internet also makes it easy for US intelligence to collect voluminous data on foreign actors.[181]

Second, many argue that China's data advantage flows from relatively lax privacy laws.[182] As Work and Grant write, "Western democracies are both wary and cautious about governments and companies massing

[174] Imbrie et al., "The Question of Comparative Advantage in Artificial Intelligence," 10.

[175] Hoadley and Sayler, "Artificial Intelligence and National Security," 30.

[176] Many observers have made this point. See Remco Zwetsloot, Helen Toner, and Jeffrey Ding, "Beyond the AI Arms Race: America, China, and the Dangers of Zero-Sum Thinking," *Foreign Affairs*, November 16, 2018, https://www.foreignaffairs.com/reviews/review-essay/2018-11-16/beyond-ai-arms-race; Carrick Flynn, "AI and data for non-experts in five minutes," *CSET*, May 2019 (unpublished); Horowitz et al., "Strategic Competition in an Era of Artificial Intelligence," 5. As one striking example, China's CloudWalk has dealt with Zimbabwe to train systems on images of black faces. Amy Hawkins, "Beijing's Big Brother Tech Needs African Faces," *Foreign Policy*, July 24, 2018, https://foreignpolicy.com/2018/07/24/beijings-big-brother-tech-needs-african-faces/.

[177] Flynn, "AI and data for non-experts in five minutes."

[178] Imbrie et al., "The Question of Comparative Advantage in Artificial Intelligence," 9.

[179] Lieber and Press, "The New Era of Counterforce," 35-42; Austin Long and Brendan Rittenhouse Green, "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategic Studies* 38 (2015): 38-73.

[180] "UCS Satellite Database," *Union of Concerned Scientists*, March 31, 2019, https://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database.

[181] Henry Farrell and Abraham L. Newman, "Weaponized Interdependence: How Global Economic Networks Shape State Coercion," *International Security* 44 (2019), 70-3; Imbrie et al., "The Question of Comparative Advantage in Artificial Intelligence," 9.

[182] Mark Zuckerberg, for example, has argued this point in Congressional testimony, urging less regulation of data use by American companies. "And the third point is — is just around enabling innovation ... we still need to make it so

personal data. Those same qualms don't exist among Chinese consumers and certainly not the Chinese government."[183] However, while China's authoritarian nature certainly means Chinese corporate data is less shielded from government eyes, this narrative oversimplifies on several accounts. Notably, Chinese consumers have increasingly demanded privacy protections from corporations, if not from the government. In part, this has been driven by concern over consumer data leaks, which one China Consumer Association found 85% of respondents had experienced.[184] When Baidu's CEO, Robin Li, remarked that "if [Chinese users] are able to exchange privacy for safety, convenience, or efficiency, in many cases they are willing to do that," widespread internet backlash erupted, with Weibo polling showing 85.8% of over 10,000 participants disagreed with Li's comments.[185] Indeed, the government has in fact stepped in to enforce data protections – in January 2019, for example, the Ministry of Industry and Information Technology blacklisted 14 mobile apps for failing to comply with personal data standards set in the May 2018 Personal Information Security Specification.[186]

In addition, American privacy protections are less robust than commonly thought. While patchwork laws exist in specific areas like healthcare, Congress has not yet passed general legislation governing the collection and use of consumer data. As such, circumstances are opposite to those in China – American citizens are much more protected from government use of data, but corporations freely exploit personal information.[187] Illustratively, the US in fact issued China at the WTO in July 2017 for overly strong privacy laws, arguing that these would disrupt normal cross-border transfers of data necessary to international commerce. The EU, whose General Data Protection Regulation (GDPR) similarly protects cross-border data transfers, declined to join the US criticism.[188]

Finally, progress on privacy-preserving machine learning may gradually reduce democratic difficulties with obtaining sufficient data to power AI applications while maintaining their values. For example, federated learning preserves privacy by never centralizing the training data into one datacenter. Rather, data remains distributed in the hands of the individuals whose data it is (e.g., within individual phones).[189] Another

---

that American companies can innovate in those areas, or else we're going to fall behind Chinese competitors and others around the world who have different regimes." See "Transcript of Mark Zuckerberg's Senate hearing," *The Washington Post*, April 8, 2018, https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/?

[183] Work and Grant, "Beating the Americans at Their Own Game," 14.

[184] Samm Sacks and Lorand Laskai, "China's Privacy Conundrum," *Slate*, February 7, 2019, https://slate.com/technology/2019/02/china-consumer-data-protection-privacy-surveillance.html.

[185] Lu Xiaomeng, Li Manyi, and Samm Sacks, "What the Facebook Scandal Means in a Land without Facebook: A Look at China's Burgeoning Data Protection Regime," *CSIS*, April 25, 2018, https://www.csis.org/analysis/what-facebook-scandal-means-land-without-facebook-look-chinas-burgeoning-data-protection.

[186] Ibid.

[187] Nuala O'Connor, "Reforming the U.S. Approach to Data Protection and Privacy," *CFR*, January 30, 2018, https://www.cfr.org/report/reforming-us-approach-data-protection.

[188] Chris Mirasola, "U.S. Criticism of China's Cybersecurity Law and the Nexus of Data Privacy and Trade Law," *Lawfare*, October 10, 2017, https://www.lawfareblog.com/us-criticism-chinas-cybersecurity-law-and-nexus-data-privacy-and-trade-law.

[189] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," *arXiv*, February 28, 2017, https://arxiv.org/pdf/1602.05629.pdf. For the most recent progress, see Keith Bonawitz, Hubert Eichner, Wolfgang

privacy-preserving machine learning method, homomorphic encryption, preserves privacy by performing machine learning only on encrypted data. Since the data owner retains the encryption keys, the ML service provider never has access to the original data. Once the provider returns the still-encrypted ML output, the data owner then uses their keys to extract the result. This would allow a medical patient, for example, to use ML on their medical records while minimizing risks of a leak.[190]

In sum, the United States appears to have a neutral to mild advantage in talent and data, and a significant advantage in hardware. There are soft reasons to expect the US talent advantage to persist, due to the tendency for talent to attract more talent, though this could be reversed by policy or circumstantial shocks; there are harder reasons, however, to expect the US hardware advantage to persist, as demonstrated in part by decades of difficulties with Chinese import substitution attempts.[191]

**(b) Leverage**

On one view, differences in inputs should end up mattering little, in terms of AI competition. If commercially-driven AI research in the United States easily spreads to China, does it matter who leads, for the purpose of military power?[192] As Joseph Nye argues, since the cost of emailing code to another is zero, the barriers to information-sharing have become dramatically lowered in the information age. We should therefore expect quick diffusion of power.[193] Brundage et al. make broadly the same point when they note the zero marginal cost of software – while building an extra tank has additional material and labor costs, duplicating a program is a simple matter of copy and pasting.[194] Unlike stealing the plans for an F-35, in this view, many AI applications are entirely based in software, and so if other countries successfully get their hands on privately developed code, they can immediately turn it to their own use.[195]

In my view, however, this is mistaken. Importantly, even if China were able to purchase exactly as much talent, hardware, and data as it desired on the open market, this would still not necessarily mean that AI's

---

Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konecny, Stefano Mazzocchi, H. Brendan McMahan, Timon Van Overveldt, David Petrou, Daniel Ramage, and Jason Roselander, "Towards Federated Learning at Scale: System Design," *arXiv*, March 22, 2019, https://arxiv.org/pdf/1902.01046.pdf.

[190] Nathan Dowlin, Ran Gilad-Bachrach, Kim Laine, Kristin Lauter, Michael Naehrig, and John Wernsing, "CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy," in *Proceedings of the 33rd International Conference on Machine Learning* (New York, NY: JMLR, 2016), 1-10. Available online: http://proceedings.mlr.press/v48/gilad-bachrach16.pdf. For a broader survey of privacy techniques and possible attacks to circumvent them, see Ho Bae, Jaehee Jang, Dahuin Jung, Hyemi Jang, Heonseok Ha, and Sungroh Yoon, "Security and Privacy Issues in Deep Learning," *arXiv*, December 6, 2018, https://arxiv.org/pdf/1807.11655.pdf.

[191] The third paper of this dissertation explores the AI supply chain in much greater detail.

[192] Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," 39.

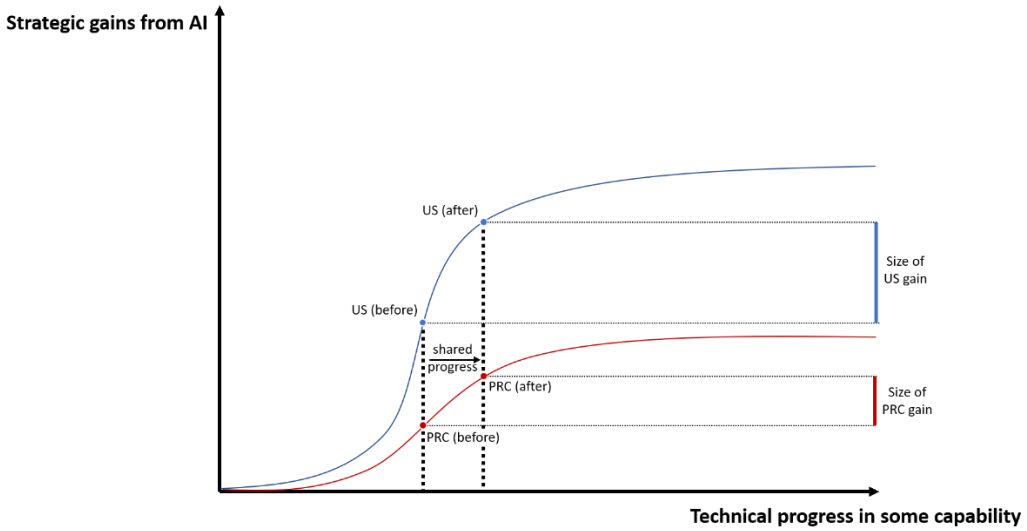
[193] Joseph S. Nye, Jr., *The Future of Power* (New York: PublicAffairs, 2011), 115.

[194] Miles Brundage, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, Paul Scharre, Thomas Zeitzoff, Bobby Filar, Hyrum Anderson, Heather Roff, Gregory C. Allen, Jacob Steinhardt, Carrick Flynn, Sean O hEigeartaigh, Simon Beard, Hadyn Belfield, Sebastian Farquhar, Clare Lyle, Rebecca Crootof, Owain Evans, Michael Page, Joanna Bryson, Roman Yapolsiky, and Dario Amodei, "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," *arXiv*, February 2018, https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf, 16-7.

[195] Hoadley and Sayler, "Artificial Intelligence and National Security," 31-2.

benefits would diffuse evenly.[196] More dramatically, even were the United States to somehow pack up Google, Facebook, and all the rest and ship the companies in a box to China, there would still be important differences in *leverage*. Here, by leverage, I mean "a state's ability to use some given technology for strategic advantage, in the context of dyadic competition with another state."

**Leverage: Technical Progress Can Play Favorites**




Let us unpack this further. In general, technologies not only demand to be acquired, in and of themselves, by would-be leading states; they also change to what degree various state characteristics, such as previously-acquired assets, surrounding geography, or operational doctrines, are still adaptive.[197] For example, Mongol light cavalry, modern navies, and ballistic missiles all changed what barriers "counted" for producing security, obviating simple remoteness, oceans, and armies still in the field, respectively, as effective shields against coercive power.[198] After the invention of motorized vehicles, horses declined rapidly as a major contributor to the military power of states. With AI, progressive substitution for human labor may analogously reduce the importance of sheer population size.[199] Conversely, some foundational shifts can also make what states are already doing, or already possess, unintentionally *more* adaptive – for

[196] Horowitz argues, applying his adoption capacity theory, that whether military-relevant AI development is public or private will drive much of its impact on relative balances of power, but this section offers an additional, important angle of analysis which may be decisive for various dyads. See Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power."

[197] Emily O. Goldman and Richard B. Andres, "Systemic effects of military innovation and diffusion," *Security Studies* 8 (1999), 116.

[198] On the Mongols, see Goldman and Andres, "Systemic effects of military innovation and diffusion," 102, 88-9. On modern power projection and the loss of American "free security," see C. Vann Woodward, "The Age of Reinterpretation," *The American Historical Review* 66 (1960): 1-19. On nuclear weapons, see Thomas C. Schelling, *Arms and Influence* (New Haven: Yale University Press, 2008), 30-1.

[199] Daron Acemoglu and Pascual Restrepo, "Demographics and Automation," *NBER*, March 2018, https://www.nber.org/papers/w24421.

example, the invention of railroads was a boon for those finding themselves naturally rich in steel precursors.[200]

Of course, if characteristics of states become maladaptive, they will change them if doing so is not more costly than the resultant benefits. But some characteristics are very costly to change, or not reasonably malleable at all as a conscious policy decision, such as regime type. In the context of dyadic competition, we might call these "strategic asymmetries." Consequently, technologies can interact with these characteristics, insofar as they differ across states, to produce enduring change in the balance of power between different states. Depending on what prior balance of military assets a state possessed before technological change, a state may thus succeed wildly in acquiring AI but still suffer a large net decrease in power due to changes in other factors.

Returning to the US-China balance of power, we can thus see how even if both sides were to acquire the same "amount" of AI, the balance might still be changed. I consider three strategic asymmetries which could mean differential ability to leverage AI for strategic gain: geography, regime type, and "enduring missions," where the last refers to particular strategic goals that the state in question has historically pursued and is unlikely to relinquish (e.g., for China, the reincorporation of Taiwan). For each, I give a number of vignettes which suggest AI's effects may fall in one direction or another; I do not intend to thoroughly analyze each vignette, but rather suggest potential strategic interactions for further analysis.

### GEOGRAPHY

Geography changes little, and thus provides an enduring source of strategic asymmetry between states.[201] For example, horse-backed cavalry made the Mongol Empire unmatched at land warfare, allowing it to rule continental Eurasia, but Japan's water barrier frustrated repeated invasion attempts, setting an eastward limit to Mongol expansion.[202] That is, Japan survived the invention of a new technology where many long-forgotten states did not because of geography. More recently, the Soviet Union's high latitude made satellites in geostationary orbit unsuited for detecting US missiles launched over the Arctic. While Soviet scientists eventually invented the highly elliptical Molniya orbit to compensate, satellites in this orbit decayed much more rapidly, requiring replacement every 1.5 years.[203] Most simply, naval advances least benefit landlocked nations.

Contextually, most plausible scenarios for US-China military conflict involve fighting in maritime East Asia. While this holds, geography's interaction with search may favor the United States, while its interaction with mass may favor China.

---

[200] Emily O. Goldman and Richard B. Andres, "Systemic effects of military innovation and diffusion," *Security Studies* 8 (1999), 116.

[201] Of course, every rule has exceptions. The Arctic is melting, generating new shipping pathways and areas of competition; we all know of China's island reclamation activities in the South China Sea. Needless to say, simple annexation (e.g., Crimea), though less common in our modern era, also changes state geographies.

[202] Goldman and Andres, "Systemic effects of military innovation and diffusion," 102-5.

[203] William Graham, "Russia's Soyuz-2-1b launches missile detection satellite," *NASASpaceFlight.com*, May 22, 2020, https://www.nasaspaceflight.com/2020/05/russias-soyuz-2-1b-missile-detection-satellite/; Yu F. Kolyuka, N. M. Ivanov, T. I. Afanasieva, and T. A. Gridchina, "Examination of the Lifetime, Evolution, and Re-Entry Features for the 'Molniya' Type Orbits" (paper presented at the 21st International Symposium of Space Flight Dynamics, Toulouse, France, 2009). Available online: https://issfd.org/ISSFD_2009/CollisionRiskII/Kolyuka.pdf.

**Search**

*Mobile Ground Targets*

On search, as discussed above, AI classifiers, which could help exploit signals and image intelligence at scale by substituting for human labor, will likely enhance states' ability to detect mobile ground targets.[204] Even if both sides acquire this capability, however, this would disproportionately favor the United States. First, AI-enabled mobile ground target detection could critically augment the US Air-Sea Battle operational concept.[205] At core, ASB aims to execute a blinding campaign against Chinese coastal assets, thereby suppressing China's ability to hold US bases and platforms at risk. A key difficulty, however, is locating China's mobile land-based missile launchers, as well as potentially dispersed and concealed command-and-control networks associated with air defense.[206]

More generally, defeating SAM mobility has become vital to modern SEAD campaigns. US efforts in Operation Desert Storm against Iraqi SAMs, which were standalone fixed targets, were over 95 percent successful; in comparison, allied efforts comprising thousands of SEAD sorties in Operation Allied Force against Serbian SAMs, which relied on mobility and concealment, destroyed only three launchers.[207] Chinese defense planners, observing this juxtaposition, increasingly emphasized mobility in the years following, with continuous investment in mobile SAM forces through the present day.[208] Critically, it is on this basis that Biddle and Oelrich (2016) argue that China will be able to maintain a coastally based A2/AD bubble indefinitely into the future, because targeting mobile ground targets against background clutter is very difficult. [209] Insofar as AI makes this less difficult, this favors the United States.

Second, mobile ground target detectability benefits China less. America's eight air bases in the Western Pacific are known, fixed targets visible on Google Maps; their usefulness does not – and could not – derive

[204] Boulanin, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk;* Geist and Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?"; Loss and Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?", *War on the Rocks*, September 19, 2019, https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/; Kallenborn, "AI Risks to Nuclear Deterrence Are Real."

[205] In 2015, AirSea Battle was renamed to Joint Concept for Access and Maneuver in the Global Commons, but is still commonly referred to by its first, more evocative name. ASB is one of several operational concepts commonly discussed as a possible US approach to conventional war with China. I do not mean to suggest ASB is the obvious operational concept the United States would actually deploy, only that by easing one of the three most-discussed approaches, this would in expectation favor the United States in the military balance. The best open-source overview is Aaron Friedberg, *Beyond Air-Sea Battle: The Debate Over US Military Strategy in Asia* (USA: Routledge, 2014).

[206] Aaron Friedberg, *Beyond Air-Sea Battle: The Debate Over US Military Strategy in Asia* (USA: Routledge, 2014), 78-84.

[207] Heginbotham et al., "The U.S.-China Military Scorecard," 127-8. In a similar vein, one oft-cited example is the "Scud hunt," where over 2,500 US Air Force sorties failed to confirm any kills against 30 Iraqi TELs. See Barry D. Watts and Thomas A. Keaney, "Effects and Effectiveness," in *Gulf War Air Power Survey, Volume II: Operations and Effects and Effectiveness*, ed. Eliot A. Cohen (Washington, DC: US Government Publishing Office, 1993), 330-40. Available online: https://apps.dtic.mil/dtic/tr/fulltext/u2/a279742.pdf. There may be other disanalogies between the Scud hunt and the Chinese case in particular. I explore this question further in the second essay. See Austin Long and Brendan Rittenhouse Green, "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategic Studies* 38 (2015), 56-64. Available online: https://www.tandfonline.com/doi/pdf/10.1080/01402390.2014.958150.

[208] Heginbotham et al., "The U.S.-China Military Scorecard," 127-8.

[209] Biddle and Oelrich, "Future Warfare in the Western Pacific," *International Security* 41.1 (2016).

from hiding.[210] Of these bases, only two are within unrefueled combat radius of the Taiwan Strait – Kadena Air Force Base and Marine Corps Air Station Futenma, both in Japan's Okinawa Prefecture.[211] Logical improvements to base survivability include hardened aircraft shelters, dispersing assets across additional bases, and shortening the runway length needed for takeoff; these tactics do not involve being elusive.[212]

*Mobile Naval Targets?*

Finally, this capability likely would not counterfactually enable Chinese detection of US surface vessels. That mission may not require modern AI at all, as detecting surface vessels is much easier than detecting mobile ground targets. Surface vessels have large minimum radar cross-sections, and distinguishing ships against the ocean is much easier than distinguishing, for example, SAMs against urban surroundings, due to the absence of background clutter. Consequently, even airborne radar can detect small cargo vessels, let alone larger ground radar or larger ships.[213] Correspondingly, US war games against China already open with US carriers steaming away from Chinese anti-ship missiles.[214]

---

[210] The eight air bases are located in Japan (Kadena AB, Futenma MCAS, Iwakuni MCAS, Yokota AB, Misawa AB), South Korea (Kunsan AB, Osan AB), and Guam (Andersen AB). On the limited number of places to be, see also John Stillion and Scott Perdue, "Air Combat Past, Present, and Future," *RAND*, August 2008, https://www.defenseindustrydaily.com/files/2008_RAND_Pacific_View_Air_Combat_Briefing.pdf, slide 67.

[211] Eric Heginbotham, Michael Nixon, Forrest E. Morgan, Jacob L. Heim, Jeff Hagen, Sheng Tao Li, Jeffrey Engstrom, Martin C. Libicki, Paul DeLuca, David A. Shlapak, David R. Frelinger, Burgess Laird, Kyle Brady, and Lyle J. Morris, "The U.S.-China Military Scorecard: Forces, Geography, and the Evolving Balance of Power, 1996–2017," *RAND*, 2015, https://www.rand.org/pubs/research_reports/RR392.html, 65.

[212] Heginbotham et al., "The U.S.-China Military Scorecard," 65-8. Fundamentally, fighting abroad means mainly depending on allied forbearance for basing. The United States cannot lead Chinese missiles on a wild goose chase across Japan; one suspects our hosts might object.

[213] Stephen Biddle and Ivan Oelrich, "Future Warfare in the Western Pacific: Technical Appendix," *Harvard Dataverse*, July 20, 2016, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/GK6PR2, 1-4; Biddle and Oelrich, "Future Warfare in the Western Pacific," *International Security* 41.1 (2016), 21-2, especially footnote 30.

[214] In particular, US carriers would seek to generate about a thousand miles of distance, placing themselves beyond the Chinese line-of-sight radar assets able to be protected from China's eastern coast. At longer ranges, China could detect US surface vessels with over-the-horizon radar, but limitations baked into physics mean OTH will remain too low-resolution to provide adequate targeting information. OTH could cue higher-resolution imaging satellites (EO, SAR) and/or airborne platforms, but the former are limited by their scarce number, implying hours of delay between useful images of a given ship, and the latter would have to enter ranges vulnerable to US countermeasures present by default for carriers. In either case, the limiting factor would not be sheer volume of intelligence, and hence AI would not solve a targeting bottleneck for China in the way I propose it might for the US above. See Christian Brose, *The Kill Chain: Defending America in the Future of High-Tech Warfare* (New York: Hachette Books, 2020), xii-xv; Heginbotham et al., "The U.S.-China Military Scorecard," 154-165; Biddle and Oelrich, "Future Warfare in the Western Pacific," especially 23; and ensuing discussion in Andrew S. Erickson, Evan Braden Montgomery, Craig Neuman, Stephen Biddle, and Ivan Oelrich, "Correspondence: How Good Are China's Antiaccess/Area-Denial Capabilities?" *International Security* 41.4 (Spring 2017): 202-13.

In short, from another angle, difficulties with detectability currently make mobile land-based assets systematically more survivable than their ocean surface equivalents. This underpins China's operational concept of "using the land to control the sea." East Asia's maritime geography thus asymmetrically benefits China; AI-enhanced intelligence-processing could reduce that asymmetry by making ground targets more detectable. See Andrew S. Erickson and David D. Yang, "Using the Land to Control the Sea? Chinese Analysts Consider the Antiship Ballistic Missile," *Naval War College Review* 62.4 (2009), 53-86.

**Mass**

*Distant Attrition*

As discussed above, one tactical effect of AI, according to various analysts, may be a shift toward swarm-on-swarm warfare. If correct, however, transitioning to this state of affairs could favor China. First, the tyranny of distance – since US-China conflict most likely occurs in East Asia, China would be sharply advantaged by proximity in any attrition-based battle, as it could reinforce losing swarms and mass forces in selected airspace with relative celerity.[215]

*Exquisite Obsolescence?*

Second, if power remains concentrated in "exquisite" platforms neither side can replace anyway, distance matters less. Thus, China arguably gains if America's exquisite platforms become obsolete, all else equal. The United States is the world leader in large, capital-intensive, high-tech platforms – China has two aircraft carriers, while we have eleven; despite their best efforts, China's fifth-generation fighter aircraft far lag the United States.[216] Many analysts have argued drone swarms in particular will threaten the continued survivability and relevance of these exquisite platforms.[217] If this is correct, even if carriers can survive massed swarms, the cost-exchange ratio of available defenses would be prohibitively unfavorable.[218] Of course, the United States could invest heavily in counter-swarm capabilities to blunt this effect.[219]

---

[215] Some suggest autonomous swarms could help Taiwan by cheapening the price of a decent air force, but even extensively substituting drones for aircraft likely would not improve the Taiwan-China air balance. Due to conventional overmatch, any airborne platforms are unlikely to be survivable beyond a few sorties, whether manned or unmanned. See Michael J. Lostumbo, David R. Frelinger, James Williams, and Barry Wilson, "Air Defense Options for Taiwan: An Assessment of Relative Costs and Operational Benefits," *RAND*, 2016, https://www.rand.org/pubs/research_reports/RR1051.html, 125-6, 70-1.

[216] Gilli and Gilli, "Why China Has Not Caught Up Yet," 178-87.

[217] There are reasons to potentially doubt this conclusion – for example, how far from land will swarms be able to operate? If the answer is "not very far," what effect do swarms have on carriers that China's missile inventory does not already have? In this case, whether an asset-based symmetry exists depends on whether those diagnosing resultant obsolescence for carriers after drone swarms are correct, and this we cannot know for certain until such capabilities are fielded. Nonetheless, I include the example for illustrative purposes, due to the frequency of its discussion.

[218] Notably, oft-envisioned drone swarms of the sort comprising thousands of platforms would necessarily require a high degree of autonomy – drones requiring human instructions would be highly vulnerable to electromagnetic jamming, and the personnel requirements of controlling thousands of drones would be prohibitive. Thus, this is not an argument against drone development in general. For reference, one day of Reaper drone operation requires 7-10 pilots, 20 sensor operators, and "scores" of intelligence analysts to sift the sensor data to boot. Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W. W. Norton and Company, 2018), 15-6.

On drone swarms and future warfare, see Scharre, "Robotics on the Battlefield Part II"; Scharre, "Robots at War and the Quality of Quantity,"; Hammes, "The Future of Warfare"; Pinion, "The Navy and Marine Corps Need to Prepare for the Swarm of the Future"; Work and Brimley, "20YY."

That all being said, this view of drone swarms is not uncontested. For a skeptical view, see Shmuel Shmuel, "The Coming Swarm Might Be Dead on Arrival," *War on the Rocks*, September 10, 2018, https://warontherocks.com/2018/09/the-coming-swarm-might-be-dead-on-arrival/. Others have suggested exquisite assets could themselves be protected with drone swarms. See Matthew Hipple, "Bring on the Countermeasure Drones," *USNI*, February 2014, https://www.usni.org/magazines/proceedings/2014/february/bring-countermeasure-drones.

[219] For proposed swarm countermeasures, see Paul Scharre, "Counter-Swarm: A Guide to Defeating Robotic Swarms," *War on the Rocks*, March 31, 2015, https://warontherocks.com/2015/03/counter-swarm-a-guide-to-defeating-robotic-swarms/; Hanacek, "The Perfect Can Wait"; Arthur Holland Michel, "Counter-Drone Systems, 2nd Edition," *Center*

Additionally, the US hardware advantage could be sufficiently large to offset suffering the obsolescence of a greater inventory of existing assets.[220]

### REGIME-TYPE

State regime types are also slow to change, and thus provide another source of strategic asymmetry. Here, AI's interactions with both search and mass may favor China. With AI, data-hungry algorithms may advantage authoritarian states, which already surveil and catalogue their own populations with little regard for human rights.[221]

## Search

*Scalable Control*

Several AI applications would likely help the Chinese Communist Party lengthen their rule.[222] First, making surveillance scalable could sharply increase the economic sustainability of intense social control.[223] Presently, autocracy requires human labor proportional to those controlled; in East Germany, for example, the Stasi employed more than 1 percent of the population.[224] China employs up to 1,000 censors per individual site, totaling two million censorship workers in 2013.[225] Since 2010, China has spent more on domestic security than external defense.[226] As a critical underlying capability, facial recognition

*for the Study of the Drone*, December 2019, https://dronecenter.bard.edu/files/2019/12/CSD-CUAS-2nd-Edition-Web.pdf.

[220] I make this argument at length in this dissertation's third essay.

[221] Dahlia Peterson, "Designing Alternatives to China's Repressive Surveillance State," *CSET*, October 2020, https://cset.georgetown.edu/research/designing-alternatives-to-chinas-repressive-surveillance-state/; Tim Hwang, "Shaping the Terrain of AI Competition," *CSET*, June 2020, https://cset.georgetown.edu/research/shaping-the-terrain-of-ai-competition/; and Andrew Imbrie, Ryan Fedasiuk, Catherine Aiken, Tarun Chhabra, and Husanjot Chahal, "Agile Alliances: How the United States and Its Allies Can Deliver a Democratic Way of AI," *CSET*, February 2020, https://cset.georgetown.edu/research/agile-alliances/.

[222] Ross Andersen, "The Panopticon Is Already Here," *The Atlantic*, 2020, https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/.

[223] Ben Angel Chang, "AI and US-China Relations," in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, ed. Nicholas Wright (DoD SMA: December 2018).

[224] Andreas Lichter, Max Loffler, and Sebastian Siegloch, "The long-term costs of government surveillance: Insights from Stasi spying in East Germany," *SOEPpapers on Multidisciplinary Panel Data Research* 865 (2016): 1-60. Available online: https://www.econstor.eu/bitstream/10419/146890/1/869045423.pdf.

[225] Beina Xu and Eleanor Albert, "Media Censorship in China," *Council on Foreign Relations*, February 17, 2017, https://www.cfr.org/backgrounder/media-censorship-china; Gary King, Jennifer Pan, and Margaret E. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *American Political Science Review* 107 (2013), 326-43.

[226] Adrian Zenz, "China's Domestic Security Spending: An Analysis of Available Data," *China Brief* 18.4 (2018), https://jamestown.org/program/chinas-domestic-security-spending-analysis-available-data/. Note that the correct interpretation of Chinese military spending data is disputed. See "What does China really spend on its military?", *China Power*, May 22, 2020, https://chinapower.csis.org/military-spending.

improvements would be vital to allow China to better exploit its 200 million physical cameras.[227] Deep learning could enable pattern recognition of when, where, and why protests are likely to break out, as well as early identification of dissident leaders, combination of various sources of data, and monitoring web platforms at scale.

*Authoritarian Economics*

Further, non-human surveillance methods may also avoid additional costs unrelated to the direct cost of employment. For example, some evidence suggests surveillance itself depresses economic activity by eroding social trust, causing individuals to reduce their productive activity.[228] Other studies have found interpersonal trust to correlate with entrepreneurship and innovation.[229] This effect was particularly acute because Stasi informants remained, publicly, their normal roles as colleagues, family, and friends, and so the knowledge of Stasi presence caused widespread doubt and fear. Automated technological surveillance would plausibly avoid these effects.[230]

Finally, search may benefit authoritarian states by enhancing state interventions into internal markets. Even given perfect information, the classic critique of centrally planning complex economies is that the computational labor required to solve an optimization problem with millions of inputs is essentially infinite, even for very high-powered supercomputers.[231] However, it is unnecessary to propose that every single economic transaction take place under the approving glance of the state to see possible benefits to relatively planned economies from AI.

Data is already used to partially plan markets, both in China and in the United States. Deep learning has already been applied to market design, as eBay, TaoBao, Amazon, Uber, and others mine their massive volumes of sales data to better match demand and supply. Several advantages obtain, including automatic pattern analysis, improved forecasting, and natural language processing for predicting demand and performing sentiment analysis. Google's "Smart Bidding," for example, uses machine learning to optimize conversions for ads; California uses AI to predict electricity demand, smoothing out the power grid and preventing blackouts.[232] Walmart's internal logistical particulars represent, to a significant degree, a centrally planned micro-economy.[233] As such, national-level strategic planning is likely to become more feasible via applying deep learning to market data.

---

[227] Paul Mozur, "One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority," *The New York Times*, April 14, 2019, https://www.nytimes.com/2019/04/14/technology/chinasurveillance-artificial-intelligence-racial-profiling.html.

[228] Ibid., 22.

[229] Stephen Knack and Philip Keefer, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," The Quarterly Journal of Economics 112 (1997): 1251-88.

[230] Lichter et al., "The long-term costs of government surveillance," 22.

[231] Ludwig von Mises, *Human Action* (Chicago: Contemporary Books, Inc., 1963), 678-80; More colloquially, see Cosma Shalizi, "In Soviet Union, Optimization Problem Solves You," *Crooked Timber*, May 30, 2012, http://crookedtimber.org/2012/05/30/insoviet-union-optimization-problem-solves-you/.

[232] Paul R. Milgrom and Steve Tadelis, "How Artificial Intelligence and Machine Learning Can Impact Market Design," forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 1-24. Available online: https://www.nber.org/books/agra-1.

[233] Leigh Phillips and Michal Rozworski, The People's Republic of Walmart: How the World's Biggest Corporations are Laying the Foundation for Socialism (Verso, 2019).

**Mass**

*Inhuman Forces*

In addition, highly capable autonomous drones would be especially useful to the CCP for suppressing unrest. Chinese counterinsurgency strategies in Tibet and Xinjiang, including ongoing mass detention and torture of Uighurs, have relied strongly on hard power solutions, as opposed to the "softer" approaches favored by the United States.[234] More generally, authoritarian states have disproportionately perpetrated mass, intentional, unrestricted killings of civilians during counterinsurgencies – a "well, just kill them all" solution to the problem of distinguishing civilians and combatants.[235] In such cases, retaining human support for the campaign has often been an important constraint; the availability of autonomous drones could remove it, thus increasingly precluding meaningful resistance in Tibet, Xinjiang, Hong Kong, and an invaded Taiwan, as well as leading to highly distasteful outcomes.[236] That is, in general, autonomous drones may reduce the cost, for authoritarian regimes, of relying on violence to maintain regime stability.[237] Asking human soldiers to kill civilians demands the regime somehow purchase a high degree of loyalty; ordering military intervention carries the risk that individual commanders may refuse to slaughter protestors, as with Major General Xu Qinxian during the Tiananman Square protests. The PLA has no parallel to DOD's Directive 3000.09, which requires that weapons systems able to select and engage targets without further human involvement allow "appropriate ... human judgment over the use of force."[238]

AI's mass effects could also resolve China's coup-proofing dilemma. Autocrats balancing external and internal threats often "coup-proof" their militaries, weakening combat power to preclude a putsch.[239] After

---

[234] Liselotte Odgaard and Thomas Galasz Nielsen, "China's Counterinsurgency Strategy in Tibet and Xinjiang," *Journal of Contemporary* China 23 (2014), 535-55. On the Uighurs, see Maya Wang, ""Eradicating Ideological Viruses": China's Campaign of Repression Against Xinjiang's Muslims," *Human Rights Watch*, September 9, 2018, https://www.hrw.org/report/2018/09/09/eradicating-ideological-viruses/chinas-campaign-repression-against-xinjiangs; Wang, "China's Algorithms of Repression"; Maya Wang, "More Evidence of China's Horrific Abuses in Xinjiang, But Little Action Holding Beijing Accountable," February 20, 2020, *Human Rights Watch*, https://www.hrw.org/news/2020/02/20/more-evidence-chinas-horrific-abuses-xinjiang.

[235] In the canonical work on this topic, Valentino et al. refer to this as a "draining the sea" strategy, as opposed to trying to catch the insurgent "fish" selectively. Benjamin Valentino, Paul Huth, and Dylan Balch-Lindsay, "'Draining the Sea': Mass Killing and Guerrilla Warfare," *International Organization* 58.2 (2004): 375-407.

[236] Valentino et al., "Draining the Sea," 396.

[237] Michael J. Boyle, "The costs and consequences of drone warfare," *International Affairs* 89 (2013), 27.

[238] On 3000.09, see Kelley M. Sayler, "Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems," CRS, December 19, 2019, https://crsreports.congress.gov/product/pdf/IF/IF11150. On the lack of a Chinese equivalent, see Elsa B. Kania, "'AI Weapons' in China's Military Innovation," *Brookings*, April 2020, https://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf, 6.

[239] Caitlin Talmadge, *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes* (Ithaca: Cornell University Press, 2015); James T. Quinlivan, "Coup-proofing: Its Practice and Consequences in the Middle East," *International Security* 24, no. 2 (1999): 131-65; Ulrich Pilster and Tobias Bohmelt, "Coup-Proofing and Military Effectiveness in Interstate Wars, 1967–99," *Conflict Management and Peace Science* 28, no. 4 (2011): 331-350; Andrew W. Bausch, "Coup-Proofing and Military Inefficiencies: An Experiment," *International Interactions* 44, no. 1 (2018): 1-32.

Formally, states face a principal-agent problem: civilian principals and military agents play a strategic game in which militaries choose to "work" or "shirk" the demands of civilians. When possible, militaries prefer to "shirk" so as to respond to the incentive structure specific to them, rather than the incentive structure faced by civilians. Authoritarian rulers fear when optimal shirking involves revolution, while democracies face only significantly mitigated risks from

Tiananmen, China incorporated its public security chiefs into core Party leadership, providing hundreds of millions of yuan in graft for loyalty.[240] As AI increasingly allows delegating power to machines rather than humans, automation would progressively reduce the total number of human beings in whose true loyalty CCP elites would need to be confident, or at least able to monitor. Consequently, China may become increasingly comfortable with decentralized command, meritocratic promotion, and other organizational practices which generate military competency, helping bridge the quality gap between Chinese and American forces.[241]

If AI substantially increases the CCP's expected lifespan, this has grand strategic implications.[242] The Soviet Union's collapse favorably resolved the Cold War without pitched battle. Besides the effect on China itself, given the CCP's habit of exporting autocratic technology, effective population control methods would risk spreading globally.[243] If more governments thus become or stay authoritarian, this would not only be normatively undesirable, but also hurt our ability to find democratic allies abroad; at the least, China could offer robust technological support to aligned dictators.[244]

---

their own armed forces, as they derive legitimacy from noncoercive sources. See Peter D. Feaver, *Armed Servants: Agency, Oversight, and Civil-Military Relations* (Cambridge: Harvard University Press, 2003), especially 8, 97, and Chapter 3; Peter Feaver, "Crisis as Shirking: An Agency Theory Explanation of the Souring of American Civil-Military Relations," *Armed Forces and Society* 24. 3 (1998): 407-34; Peter Feaver, "The Civil-Military Problematique: Huntington, Janowitz, and the Question of Civilian Control," *Armed Forces and Society* 23.2 (1996): 149-78. See also David Pion-Berlin and Harold Trinkunas, "Civilian Praetorianism and Military Shirking During Constitutional Crises in Latin America," *Comparative Politics* 42, no. 4 (2010): 395-411. On democracies, see Talmadge, *The Dictator's Army*, 3, 19-23.

[240] Yuhua Wang and Carl Minzner, "The Rise of the Chinese Security State," *The China Quarterly* 222 (2015): 339-59; Yuhua Wang, "Empowering the Police: How the Chinese Communist Party Manages Its Coercive Leaders," *The China Quarterly* 219 (2014): 625-48; Yuhua Wang, "Coercive capacity and the durability of the Chinese communist state," *Communist and Post-Communist Studies* 47 (2014): 13-25.

[241] Talmadge, *The Dictator's Army*; Steven Feldstein, "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression," *Journal of Democracy* 30 (2019), 42.

[242] For the argument the CCP will collapse, see David Shambaugh, "The Coming Chinese Crackup," *The Wall Street Journal*, March 6, 2015, https://www.wsj.com/articles/the-coming-chinese-crack-up-1425659198.

Other scholars argue a democratic China would become a responsible state, lessening the threat to the United States sufficiently to allow withdrawal from East Asia. At the least, the risk of great power war might be reduced. See Aaron Friedberg, *A Contest for Supremacy: China, America, and the Struggle for Mastery in Asia* (New York: W. W. Norton & Company, 2011), 250-2, 264-84; Daniel M. Kliman, *Fateful Transitions: How Democracies Manage Rising Powers, from the Eve of World War I to China's Ascendance* (Philadelphia: University of Pennsylvania Press, 2015).

[243] Andrew J. Nathan, "China's Challenge," *Journal of Democracy* 26.1 (2015): 156-170. Even absent extensive Chinese exports, other autocratic countries would likely copy successful methods. See Tim Hwang, "Shaping the Terrain of AI Competition," *CSET*, June 2020, https://cset.georgetown.edu/research/shaping-the-terrain-of-ai-competition/, 19.

[244] Increased ideological distance between ourselves and potential future allies likely makes coordination more difficult. Democracies more easily ally with other democracies, although there are notable exceptions. Mark L. Haas, *The Ideological Origins of Great Power Politics, 1789-1989* (Ithaca: Cornell University Press, 2005); Bruce Russett, *Grasping the Democratic Peace: Principles for a Post–Cold War World* (New Jersey: Princeton University Press, 1993); Bruce Russett and John R. Oneal, *Triangulating Peace: Democracy, Interdependence, and International Organizations* (New York: Norton, 2001); Michael W. Doyle, *Liberal Peace: Selected Essays* (New York: Routledge, 2011).

*Information Warfare*

Finally, besides increasing the fitness of authoritarian governments more generally, highly effective, AI-enhanced information warfare may unlock persuasion as a viable theory of victory for governments willing to deploy manipulative propaganda at scale. If mass opinion comes to be decisively influenced by the clash between AI influence systems, for example, China may come to believe that its best bet for reabsorbing Taiwan is heavy investment in AI-empowered propaganda. After all, China's Great Firewall arguably represents a ready-made testbed for future AI-enabled information control techniques.

This would more generally pose a significant challenge for democratic societies, as precision propaganda would increase the efficacy of forms of political interference differentially used more frequently by authoritarian states. Russian interference into recent US presidential elections was accomplished with only modern-day systems. As AI capabilities develop, it may become harder and harder for democratic states to meaningfully aggregate the opinions of their own populations, insofar as an increasingly greater percentage of the information available to those populations becomes manufactured. Leveraging AI to alter target states' national priorities through information warfare would represent "winning without fighting" *par excellence*.[245]

## ENDURING MISSIONS

Finally, for various reasons, states tend to adopt enduring missions which are unlikely to change over time, representing a third and final source of strategic asymmetries. For China, reincorporating Taiwan is central to the CCP's legitimacy. For the United States, flirting with nuclear counterforce has seemingly always been in our DNA.[246] Search and mass variously change the difficulty of these missions, thereby affecting the balance of power by potentially changing how relevant scenarios would resolve.

### Search

*Occupation Policing*

First, given the order-of-magnitude difference between the two governments' military budgets, Taiwanese victory may increasingly require protracted guerilla warfare in the island's mountains and cities, as opposed to preventing a Chinese amphibious landing outright.[247] Even assuming US intent to intervene, China's

---

On AI driving ideological bifurcation among states, see Nicholas Wright, "How Artificial Intelligence Will Reshape the Global Order," *Foreign Affairs*, July 10, 2018, https://www.foreignaffairs.com/articles/world/2018-07-10/howartificial-intelligence-will-reshape-global-order; Chang, "AI and US-China Relations."

[245] Richard Danzig, "An Irresistible Force Meets a Moveable Object: The Technology Tsunami and the Liberal Order," *Lawfare Research Paper Series* 5.1 (2017), https://assets.documentcloud.org/documents/3982439/Danzig-LRPS1.pdf, 4-7.

[246] Francis Gavin, Nuclear Statecraft: History and Strategy in America's Atomic Age (US: Cornell University, 2012); Austin Long and Brendan Rittenhouse Green, "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategic Studies* 38 (2015): 38-73.

[247] Analysts disagree whether China today could successfully land forces amphibiously, though all agree trend lines do not favor Taiwan. Previously, the "porcupine" strategy sought to deny a Chinese landing, such as by developing Taiwanese A2/AD capabilities. This operational concept is still strategically necessary if one believes Taiwan's citizens lack political will, such that any successful amphibious landing by the PRC would prompt quick surrender.

In contrast, the "hard ROC" strategy seeks to guarantee a long, exhausting, costly struggle for the PRC after landing on Taiwan, such as by developing layered ground defenses and preparing for extended guerilla warfare. This operational concept is viable if Taiwan's citizens are willing to accept protracted war in defense of their freedoms.

missile inventory could mean US base closures for days to weeks at war's start, requiring meaningful Taiwanese resistance to stave off a *fait accompli*.[248] As Taiwan's citizens increasingly identify as "just Taiwanese" (66 percent, according to 2019 Pew polling), as opposed to "both Taiwanese and Chinese" (28 percent) or "just Chinese" (4 percent), holding out may be costly but plausible.[249] Analogously, during the 1979 Sino-Vietnamese war, a conventionally inferior Vietnam nonetheless denied China a clear victory.[250]

As population-wide resistance becomes increasingly critical to Taiwan's defense, AI-enabled capabilities which make Taiwan more digestible would thus strike at the core of its deterrent. With search, predictive policing could help China distinguish Taiwanese combatants from neutral civilians. Data from social media, visual surveillance, and electronic eavesdropping could be automatically processed to identify guerillas without mass interrogation, reducing collateral damage and building local support.[251]

---

The defining work on the "porcupine" strategy is William S. Murray, "Revisiting Taiwan's Defense Strategy," *Naval War College Review* 61.3 (2008), https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1814&context=nwc-review. The defining work on the "hard ROC" strategy is Jim Thomas, John Stillion, and Iskander Rehman, "Hard ROC 2.0: Taiwan and Deterrence through Protraction," *CSBA*, 2014, https://csbaonline.org/uploads/documents/2014-10-01_CSBA-TaiwanReport-1.pdf. Most concisely, see p. 57 for an informative diagram. As the report notes, this would be a "strategy of Fabian defense ... Taiwanese ground forces could melt into the island's urban and mountainous areas in order to wage a 'war of a thousand cuts' against PLA occupation forces" (vi). Usefully condensed discussion of these and other options is available at Jennifer M. Turner, "The Cost of Credible Deterrence in Taiwan," *War on the Rocks*, January 13, 2016, https://warontherocks.com/2016/01/the-cost-of-credible-deterrence-in-taiwan/; and at James R. Holmes, "Taiwan Needs a Maoist Military," *Foreign Policy*, October 17, 2019, https://foreignpolicy.com/2019/10/17/taiwan-maoist-military-china-navy-south-china-sea/.

On trend lines in Taiwanese ability to resist amphibious invasion, see "Military and Security Developments Involving the People's Republic of China 2019: Annual Report to Congress:," *Office of the Secretary of Defense*, May 2, 2019, https://media.defense.gov/2019/May/02/2002127082/-1/-1/1/2019_CHINA_MILITARY_POWER_REPORT.pdf, 90-1, 85; see also discussion at Scott L. Kastner, "Is the Taiwan Strait Still a Flash Point? Rethinking the Prospects for Armed Conflict between China and Taiwan," *International Security* 40.3 (2015/16), 69-72; and David A. Shlapak, David T. Orletsky, Toy I. Reid, Murray Scot Tanner, Barry Wilson, "A Question of Balance: Political Context and Military Aspects of the China-Taiwan Dispute," *RAND*, 2009, https://www.rand.org/pubs/monographs/MG888.html, especially 123, 139-40, but also 91-121.

[248] Eric Heginbotham, Michael Nixon, Forrest E. Morgan, Jacob L. Heim, Jeff Hagen, Sheng Tao Li, Jeffrey Engstrom, Martin C. Libicki, Paul DeLuca, David A. Shlapak, David R. Frelinger, Burgess Laird, Kyle Brady, and Lyle J. Morris, "The U.S.-China Military Scorecard: Forces, Geography, and the Evolving Balance of Power, 1996–2017," *RAND*, 2015, https://www.rand.org/pubs/research_reports/RR392.html, xxiii, 45-68. See also Thomas Shugart and Javier Gonzalez, "First Strike: China's Missile Threat to U.S. Bases in Asia," *CNAS*, June 2017, https://s3.amazonaws.com/files.cnas.org/documents/CNASReport-FirstStrike-Final.pdf?.

[249] Kat Devlin and Christine Huang, "In Taiwan, Views of Mainland China Mostly Negative," *Pew Research*, May 12, 2020, https://www.pewresearch.org/global/2020/05/12/in-taiwan-views-of-mainland-china-mostly-negative/. See also Kastner, "Is the Taiwan Strait Still a Flash Point?," 76; Ian Easton, *The Chinese Invasion Threat: Taiwan's Defense and American Strategy in Asia* (US: Eastbridge Brooks, 2019), 193.

[250] Thomas et al., "Hard ROC 2.0," 60.

[251] For a primer, see Tim Lau, "Predictive Policing Explained," *Brennan Center for Justice*, April 1, 2020, https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained. For a long-form introduction, see Walter L. Perry, Brian McInnis, Carter C. Price, Susan C. Smith, John S. Hollywood, "Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations," *RAND*, 2013, https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf.

*Nuclear Counterforce*

Second, conversely, in the nuclear domain, locating China's road-mobile missiles is the key challenge limiting US counterforce, executed either as part of damage-limitation after the outbreak of nuclear war or, less likely, as a disarming first strike.[252] Timely identification is primarily complicated by "needle in a haystack" dynamics – China is large, and analytic labor is scarce.[253] As with elusive ground targets, AI-enabled intelligence processing of satellite imagery, unattended ground sensors, drone imagery, and other intelligence streams could provide the necessary scale and speed, as I explore at paper length below.[254] Conversely, China simply does not have enough nuclear weapons for meaningful counterforce, even knowing the location of every American asset.[255]

More generally, China's posture of assured retaliation accepts nuclear vulnerability but promises revenge, while our damage-limitation strategy seeks to reduce vulnerability to begin with.[256] Our strategy comprises two parts: a disarming first-strike against China's arsenal, and absorption of surviving weapons with missile defense.[257] Some analysts argue AI could help missile defenses distinguish decoys from real warheads by training reliable classifiers; since China has no strategic missile defense capability, this would only benefit the United States.[258] Of course, it bears repeating that asset-based asymmetries most easily erode; here, several possible future Chinese acquisitions, such as nuclear-tipped hypersonic glide vehicles, competently quiet SSBNs, or simply a much larger arsenal, could change this calculus.

---

[252] Wu Riqiang, "Living with Uncertainty: Modeling China's Nuclear Survivability," *International Security* 44.4 (2020), 92-3; Heginbotham et al., "The U.S.-China Military Scorecard," 310-2. I explore the question of AI-assisted nuclear counterforce at length in this dissertation's second essay.

[253] Harper, "Artificial Intelligence to Sort Through ISR Data Glut"; Cardillo, "Small Satellites – Big Data."

[254] Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*; Geist and Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?"; Loss and Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?"; Johnson, "MAD in an AI Future?"; Loss, "Artificial Intelligence, the Final Piece to the Counterforce Puzzle?"; Kallenborn, "AI Risks to Nuclear Deterrence Are Real"; Lieber and Press, "The New Era of Counterforce."

[255] The Chinese arsenal is simply of insufficient size. In 2015, modeling by RAND assessed that even assuming various unknowns favor China, including a high-end Chinese arsenal estimate, US arsenal reductions due to New START, and non-functional US missile defense, a disarming first strike in 2017 would still leave at least 988 US weapons intact. See Heginbotham et al., "The U.S.-China Military Scorecard," 312-4.

[256] Fiona S. Cunningham and M. Taylor Fravel, "Assuring Assured Retaliation: China's Nuclear Posture and U.S.-China Strategic Stability," *International Security* 40.2 (2015): 7-50.

[257] Charles L. Glaser and Steve Fetter, "Should the United States Reject MAD? Damage Limitation and U.S. Nuclear Strategy toward China," *International Security* 41.1 (2016): 49-98; and ensuing discussion at Brendan Rittenhouse Green, Austin Long, Matthew Kroenig, Charles L. Glaser, and Steve Fetter, "Correspondence: The Limits of Damage Limitation," *International Security* 42.1 (2017): 193-207.

[258] Vincent Boulanin, "The future of machine learning and autonomy in nuclear weapon systems," in *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume I: Euro-Atlantic Perspectives*, ed. Vincent Boulanin, SIPRI, May 2019, https://www.sipri.org/sites/default/files/2019-05/sipri1905-ai-strategic-stability-nuclear-risk.pdf, 54-5, 58-9; Zachary S. Davis, "Artificial Intelligence on the Battlefield: An Initial Survey of Potential Implications for Deterrence, Stability, and Strategic Surprise," *Center for Global Security Research*, March 2019, https://cgsr.llnl.gov/content/assets/docs/CGSR-AI_BattlefieldWEB.pdf, 6-8, 14.

**Mass**

*Urban Warfare*

AI's mass effects could also help China take Taiwan by strengthening urban warfare capabilities. Automating the most dangerous tasks could reduce PLA casualties, helping maintain domestic support for a long war. As above, AI is likely to ease the difficulties of urban warfare through various applications of autonomous platforms.

*Political Warfare*

Further, AI could enhance Chinese political warfare. The CCP has long wielded disinformation against Taiwan, including through "fake news" campaigns supporting preferred candidates in the island's 2020 elections.[259] As computer-generated writing, video, and audio become increasingly indistinguishable from reality, voter access to true information may degrade, hurting the Taiwanese government's ability to reflect the preferences of its citizenry.[260] AI could also improve Chinese manipulation of those preferences outright – by processing social media posts, demographic information, and other data sources, algorithms could tailor persuasion to each individual's personality type, social connections, and political beliefs.[261]

---

[259] According to a Swedish think tank studying foreign disinformation, Taiwan ranked first in its level of exposure. See Alice Su, "Can fact-checkers save Taiwan from a flood of Chinese fake news?," *Los Angeles Times*, December 16, 2019, https://www.latimes.com/world-nation/story/2019-12-16/taiwan-the-new-frontier-of-disinformation-battles-chinese-fake-news-as-elections-approach; Lily Kuo, "Taiwan's citizens battle pro-China fake news campaigns as election nears," *The Guardian*, December 30, 2019; Chia-Chien Chang and Alan H. Yang, "Weaponized Interdependence: China's Economic Statecraft and Social Penetration Against Taiwan," *Orbis* 64.2 (2020): 312-33. Notably, for concern that ongoing Taiwanese efforts to resist fake news may risk freedom of speech, see Nick Aspinwall, "Taiwan's War on Fake News Is Hitting the Wrong Targets," *Foreign Policy*, January 10, 2020, foreignpolicy.com/2020/01/10/taiwan-election-tsai-disinformation-china-war-fake-news-hitting-wrong-targets/.

[260] Robert Chesney and Danielle K. Citron, "Disinformation on Steroids," *CFR*, October 16, 2018, https://www.cfr.org/report/deep-fake-disinformation-steroids; Joe Littell, "Don't Believe Your Eyes (or Ears): The Weaponization of Artificial Intelligence, Machine Learning, and Deepfakes," War on the Rocks, October 7, 2019. Surveying 500 respondents, Kreps and McCain found that 72 percent found a GPT-2 generated New York Times article about North Korea credible, compared to 83 percent for a real Times piece. See Kreps and McCain, "Not Your Father's Bots."

Two AI-enabled disinformation technologies are deepfakes and computer-generated fake news. Deepfakes are AI-created images, video, and audio which impersonate real human beings, and which are increasingly difficult to distinguish from genuine media. Computer-generated fake news, sometimes called "neural fake news," consists of AI-created text increasingly difficult to distinguish from human writing. By alleviating human labor constraints, both techniques are thought to potentially enable disinformation at scale. On deepfakes, see a recent survey at Yisroel Mirsky and Wenke Lee, "The Creation and Detection of Deepfakes: A Survey," *arXiv*, May 12, 2020, https://arxiv.org/pdf/2004.11138.pdf. On computer-generated fake news, see analysis of the most well-known generator, GPT-2, at Sarah Kreps and Miles McCain, "Not Your Father's Bots: AI is Making Fake News Look Real," *Foreign Affairs*, August 2, 2019, www.foreignaffairs.com/articles/2019-08-02/not-your-fathers-bots. See reflection by OpenAI itself on its phased release strategy at Irene Solaiman, Miles Brundage, Jack Clark, Amanda Askell, Ariel Herbert-Voss, Jeff Wu, Alec Radford, Gretchen Krueger, Jong Wook Kim, Sarah Kreps, Miles McCain, Alex Newhouse, Jason Blazakis, Kris McGuffie, and Jasmine Wang, "Release Strategies and the Social Impacts of Language Models," *arXiv*, November 13, 2019, https://arxiv.org/ftp/arxiv/papers/1908/1908.09203.pdf. See discussion of technical solutions at Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi, "Defending Against Neural Fake News," *arXiv*, October 29, 2019, https://arxiv.org/pdf/1905.12616.pdf. For general discussion of other AI-enabled mass-manipulation capabilities, see Matt Chessen, "The MADCOM future," *Atlantic Council*, 2017, www.atlanticcouncil.org/in-depth-research-reports/report/the-madcom-future/.

[261] Dipayan Ghosh and Ben Scott, "Digital Deceit: The Technologies Behind Precision Propaganda on the Internet," *New America*, January 2018, https://d1y8sb8igg2f8e.cloudfront.net/documents/digital-deceit-final-v3.pdf, 26-30;

## Conclusion

This paper sought to help tame the nascent, but already highly ensnarled, discussion of AI by security studies scholars. First, I argued that security studies should focus its study of AI on "deep learning," the modern AI technique driving most recent headlining progress. Without focus, "AI" by default ranges tax software, drones, thermostats, ballistic missile defenses, brute-force computation, machine learning, and all manner of other grab-bag vagaries. If we wish to answer questions like, "how will AI affect the US-China balance?", we will require a suitably sized independent variable.

Second, leaning on my narrower operationalization, I reviewed the computer science literature about what deep learning *does*, and conceptualized its effects in a manner hopefully more suitable for security studies. I argued that AI's mechanical effects can be summarized as "pattern recognition and pattern generation, minus misalignment," and that its tactical effects can similarly be encapsulated as "search plus mass minus slippage."

Finally, a significant literature asks whether China might, exploiting AI, leapfrog US power. In the last part of this chapter, I unpacked this inquiry into two questions about US and Chinese state-specific factors: first, does China have more of the inputs – i.e., talent, hardware, and data – which serve as key precursors to AI? Second, even if the US or China successfully translates plentiful resources into many deployed military applications of AI, how useful are those applications ("leverage") in terms of the balance of power between the United States and China? Although the assessments in this paper are preliminary, I found that AI may favor the US in the nuclear domain, assist China with peacetime competition, and produce mixed effects in the conventional domain.

**AI's Leverage-Based Effects**

| | Nuclear | Conventional | Peacetime |
|---|---|---|---|
| **US-Favoring** | Easier Counterforce | Easier Air-Sea Battle | --- |
| **China-Favoring** | --- | Easier Taiwan;<br>Easier Air-to-Air | Easier Authoritarianism:<br>Use of Force, Economics |

If AI's effects on international affairs seem, even at this early stage, likely to be sufficiently large, it behooves the discipline of security studies to begin now grappling with its implications, if only to advise policymakers already making decisions with long-run impact in the present. Despite the world's militaries not having yet fielded mature AI-enabled weapons systems, a variety of research methodologies represent available avenues for scholarly progress.[262]

> ● **First, theory.** International relations comprises a rich population of causal models about state behavior, technology, and power, and it seems likely that significant low-hanging fruit exists in applying that existing literature to AI's particular features. Since applying theory in turn requires technically sound excavation of what AI's particular features *are*, in a way which makes the technology "legible" to the security studies discipline's theoretical constructs, work in this bucket includes both conceptualizing AI's effects and adapting extant theories to those effects. This paper is a first-cut effort in this category.

Chessen, "The MADCOM future"; Frank Adkins and Shawn Hibbard, "The Coming Automation of Propaganda," *War on the Rocks*, August 6, 2019, https://warontherocks.com/2019/08/the-coming-automation-of-propaganda/.

[262] Michael Horowitz, "Do Emerging Military Technologies Matter for International Politics?", *Annual Review of Political Science* 23 (2020), 387.

● **Second, surveys.** Novel data can be produced through eliciting both mass and expert opinions.[263] For those seeking to assess whether, when, and why AI capabilities may cross theoretically important thresholds, polling actual AI researchers can generate predictive data.[264] Understanding the attitudes of likely military end-users for AI systems, including to what degree they welcome, trust, or fear battlefield use, can inform theory about AI's impact on military conflict.[265]

● **Third, wargames.** A surge of recent scholarship has explored wargaming as a generalizable experimental method for political science research.[266] Since wargames generate data, albeit data produced by a context which must be carefully considered for disanalogies to actual warfare, they are especially useful when history contains little or no examples of the phenomenon studied, such as extended drone wars or mutual nuclear exchange.[267] Consequently, wargaming methods seem especially suitable for application to scenarios involving AI-enabled weapons systems.[268]

● **Fourth, formal modeling.** AI-integrated militaries may raise questions about credible commitments, signaling, and delegation which seem especially amenable to game-theoretic formal modeling. Fully autonomous weapons systems, free from human cognitive biases, may uniquely conform to rationality assumptions.[269] In recent work, Garfinkel and Dafoe (2019) begin to reason mathematically about how swarm warfare could affect the offense-defense balance.[270] These methods could be applied to reason about AI across different battlefield conditions.[271]

---

[263] Baobao Zhang and Allan Dafoe, "U.S. Public Opinion on the Governance of Artificial Intelligence," Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society, 2020; Baobao Zhang and Allan Dafoe, "Artificial Intelligence: American Attitudes and Trends," Centre for the Governance of AI, 2019; Baobao Zhang, Markus Anderljung, Lauren Kahn, Noemi Dreksler, Michael C. Horowitz, and Allan Dafoe, "Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers," Center for the Governance of AI, 2019.

[264] Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans, "When Will AI Exceed Human Performance? Evidence from AI Experts," *Journal of Artificial Intelligence Research* 62 (2018), 729-54.

[265] Macdonald and Schneider, "Battlefield Responses to New Technologies," 216-49.

[266] Erik Lin-Greenberg, Reid Pauly, and Jacquelyn Schneider, "Wargaming for Political Science Research," *SSRN*, 2021, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3676665.

[267] Erik Lin-Greenberg, "Remote Controlled Restraint: The Effect of Remote Warfighting Technology on Crisis Escalation" (PhD diss., Columbia University, 2019); Reid Pauly, "Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint," *International Security* 43.2 (2018): 151-92; Jacquelyn Schneider, "Cyber and Crisis Escalation: Insights from Wargaming," *USASOC Futures Forum*, 2017.

[268] Yuna Huh Wong, John Yurchak, Robert W. Button, Aaron Frank, Burgess Laird, Osonde A. Osoba, Randall Steeb, Benjamin N. Harris, and Sebastian Joon Bae, "Deterrence in the Age of Thinking Machines," *RAND*, 2020, https://www.rand.org/pubs/research_reports/RR2797.html.

[269] Kenneth Payne, *Strategy, Evolution, and War: From Apes to Artificial Intelligence* (US: Georgetown University Press, 2018).

[270] Ben Garfinkel and Allan Dafoe, "How does the offense-defense balance scale?," *Journal of Strategic Studies* 42.6 (2019): 736-63.

[271] A recent, thorough effort can be found at Stephen Biddle, *Military Power: Explaining Victory and Defeat in Modern Battle* (Princeton: Princeton University Press, 2004), 209-40.

● **Finally, campaign analysis.** The use of campaign analysis by security studies dates back at least to the Cold War, when John Mearsheimer and Barry Posen debated the conventional balance in Central Europe in the pages of *International Security*.[272] Extant work by Heginbotham et al. (2015) and Lieber and Press (2017) already touch, however lightly, on AI applications affecting the US-China military balance, and recent scholarship formalizing the campaign analysis method should ease its further application.[273]

This paper sought to cover significant conceptual ground, in seeking to make AI legible to security studies as a construct amenable to study. In doing so, it doubtless generated some number of errors. Given the importance of AI to our discipline, it is hoped that the aggravated reader will take such errors, if they are also convinced of AI's importance, as a provocation to join a sorely needed discussion.

[272] John Mearsheimer, "Why the Soviets Can't Win Quickly in Central Europe," *International Security* 7.1 (1982), 3-39; Barry Posen, "Measuring the European Conventional Balance: Coping with Complexity in Threat Assessment," *International Security* 9.3 (1984), 47–88.

[273] Lieber and Press, "The New Era of Counterforce"; Heginbotham et al., "The U.S.-China Military Scorecard"; Rachel Tecott and Andrew Halterman, "The Case for Campaign Analysis: A Method for Studying Military Operations," *SSRN*, 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3688529.

# Artificially Assured Destruction?
# Modeling the Effects of Artificial Intelligence on the US-China Nuclear Balance

*with Torin Rudeen*

## Introduction

Will artificial intelligence endanger China's second-strike capability against the United States? If AI means a first-strike counterforce effort by the United States would become overwhelmingly likely to succeed, decision-makers in both the United States and China would face grim dilemmas during an escalating crisis. Washington, afraid of losing the chance to "win cleanly and quickly" and avoid a long, costly conflict with an uncertain outcome, would be tempted to decide matters in one nuclear swoop while it still could. Conversely, Beijing, fearing this calculus, would face "use it or lose it" pressures for first use.[274] In short, under first-strike instability, to paraphrase the Cold War adage, the only thing worse than going first is going second.[275] Thus, whether or not AI may induce first-strike instability in the US-China dyad specifically, and in other "suspicious" dyads like India and Pakistan more generally, is of critical importance.

More generally, if AI strongly enhances leading militaries' ability to conduct nuclear counterforce, this may implicate the sustainability of unipolarity most generally. If the nuclear revolution holds, continued great power rivalry arguably seems mysterious for all but the most obstinate flavors of offensive realism, as states can secure absolute security for themselves through acquiring a survivable arsenal.[276] Further, since the world economy is relatively open in historical terms, many non-survival interests related to flourishing can also theoretically be obtained without some spasm of revisionist conquest.[277] Consequently, waging war in

[274] Avery Goldstein, "First Things First: The Pressing Danger of Crisis Instability in U.S.-China Relations," *International Security* 37.4 (2013): 49-89; Caitlin Talmadge, "Would China Go Nuclear? Assessing the Risk of Chinese Nuclear Escalation in a Conventional War with the United States," *International Security* 41, no. 4 (2017): 50-92; Joshua Rovner, "Two kinds of catastrophe: nuclear escalation and protracted war in Asia," Journal of Strategic Studies 40, no. 5 (2017): 696-730; David C. Gompert, Astrid Stuth Cevallos, and Cristina L. Garafola, "War with China: Thinking Through the Unthinkable," *RAND*, 2016, https://www.rand.org/pubs/research_reports/RR1140.html; Eric Heginbotham, Michael S. Chase, Jacob L. Heim, Bonny Lin, Mark R. Cozad, Lyle J. Morris, Christopher P. Twomey, Forrest E. Morgan, Michael Nixon, Cristina L. Garafola, and Samuel K. Berkowitz, "China's Evolving Nuclear Deterrent: Major Drivers and Issues for the United States," *RAND*, 2017, https://www.rand.org/pubs/research_reports/RR1628.html; Rebecca Davis Gibbons and Matthew Kroenig, "Reconceptualizing nuclear risks: Bringing deliberate nuclear use back in," *Comparative Strategy* 35, no. 5 (2016): 407-22; Michael S. Chase, Andrew S. Erickson, and Christopher Yeaw, "Chinese Theater and Strategic Missile Force Modernization and its Implications for the United States," *Journal of Strategic Studies* 32, no. 1 (2009): 67-114; Charles L. Glaser and Steve Fetter, "Should the United States Reject MAD? Damage Limitation and U.S. Nuclear Strategy toward China," *International Security* 41, no. 1 (2016): 49-98; Brendan Rittenhouse, Austin Long, Matthew Kroenig, Charles L. Glaser, and Steve Fetter, "Correspondence: The Limits of Damage Limitation," *International Security* 42 (2017), 193-207.

[275] Goldstein, "First Things First," 88. First-strike stability holds "when, after, considering the vulnerability of strategic forces on both sides, neither leader perceives the other as pressured by the posture of forces to strike first in a crisis." See Glenn A. Kent and David E. Thaler, "First-Strike Stability: A Methodology for Evaluating Strategic Forces," *RAND*, 1989, https://www.rand.org/pubs/reports/R3765.html.

[276] Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (New York: Cornell University Press, 1989); ---, *The Illogic of American Nuclear Strategy* (New York: Cornell University Press, 1984).

[277] G. John Ikenberry, *A World Safe for Democracy: Liberal Internationalism and the Crises of Global Order* (US: Yale University Press, 2020); Michael Mousseau, "The End of War: How a Robust Marketplace and Liberal Hegemony Are Leading to Perpetual World Peace," *International Security* 44.1 (2019): 160-96.

a risky bid to seize hegemony seems, given a nuclear arsenal's guarantee against invasion, perhaps the only way a regional power can place its survival in real doubt at all.[278] Most prominently, Nuno Monteiro's theory of unipolar politics synthesizes these conditions by arguing that if the nuclear revolution holds (and war is thus costly), then unipolar powers can indefinitely enjoy a durable world order in which they are *primus inter pares* by accommodating other powers' economic needs.[279]

Another thread in international relations theory, however, argues that orders inherently generate disequilibria over time, leading to the punctuation of world orders with wars which settle into the establishment of new systems.[280] Technology figures prominently as a causal intermediary in these sorts of cyclic views, often playing the role of both endogenous instrument of would-be disruptors as well as exogenous disruptor itself.[281] If both hegemons and rising powers can both enjoy security and growth, after all, what is there to fight over? Here, AI serves as an example of exactly what – diffusion of specifically dual-use AI applications seems almost inevitable alongside continued global economic growth, after all, due to their economically profitable nature.[282] Such dual-use applications, however – in particular, order-of-magnitude jumps in the scale and speed of data-processing – exactly underlie the causal mechanisms our model finds may endanger China's second-strike capability.

Consequently, if such technologies and their successors naturally make nuclear counterforce more thinkable for the unipole, guaranteeing nuclear survivability (and signaling credible intentions to do as much) becomes difficult for even benevolently intentioned hegemons seeking to pacify would-be challengers. Thus, for theorists seeking to square the circle of mature, long-held nuclear arsenals, deep economic interdependence following decades of globalization, and continued US-China rivalry, the structural inevitability of counterforce-enabling technologies underpinned by AI provides one causal explanation for why US-China competition has progressed toward ever fiercer stages: despite all our wealth and strength, the continued march of technology simply means states are not yet safe. Competition hedges against technological development, not only in the nuclear realm, but in the realm of grand strategic choice between satisfied behavior as a status quo power and revisionism more generally.[283]

---

[278] Jonathan Kirshner, "The tragedy of offensive realism: Classical realism and the rise of China," *European Journal of International Relations* 18.1 (2012): 53-75.

[279] Nuno Monteiro, *Theory of Unipolar Politics* (New York: Cambridge University Press, 2014); an excellent, lively debate is available at Nuno Monteiro, William C. Wohlforth, Michael Beckley, Christopher Layne, and Jeffrey W. Taliaferro, "Theory of Unipolar Politics," *Roundtable* 8.3 (2015), https://issforum.org/ISSF/PDF/ISSF-Roundtable-8-3.pdf.

[280] Robert Gilpin, *War and Change in World Politics* (UK: Cambridge University Press, 1981); Paul Kennedy, *The Rise and Fall of the Great Powers* (New York: Random House, 1987).

[281] George Modelski and William R. Thompson, *Leading Sectors and World Powers* (Columbia: University of South Carolina Press, 1996); Daniel W. Drezner, "Technological change and international relations," *International Relations* 33.2 (2019); Allan Dafoe, "On Technological Determinism: A Typology, Scope Conditions, and a Mechanism," *Science, Technology, and Human Values* 40.6 (2015), 1047–76; Emily O. Goldman and Richard B. Andres, "Systemic effects of military innovation and diffusion," *Security Studies* 8 (1999): 79-125; Henry Kissinger, *World Order* (USA: Penguin Books Limited, 2014)

[282] Michael C. Horowitz, *The Diffusion of Military Power: Causes and Consequences for International Politics* (New Jersey: Princeton University Press, 2010).

[283] Brendan Rittenhouse Green, *The Revolution That Failed: Nuclear Competition, Arms Control, and the Cold War* (UK: Cambridge University Press, 2020); Keir Lieber and Daryl Press, *The Myth of the Nuclear Revolution: Power Politics in the Atomic Age* (New York: Cornell University, 2020).

Predictably, these substantive and theoretical implications have provoked an increasingly vast literature assessing AI's impact on the US-China nuclear balance.[284] At present, however, this literature inspires two dissatisfactions: first, lacking a consensus definition of AI, the field discusses an enormous panoply of commonsensically distinct technologies under one banner. Different works appear to reference technologies falling under different operationalizations of AI, each of which carry their own technical limits, vulnerabilities, and capabilities; to some extent, scholars are talking past one another.[285] Second, since AI's adoption by modern militaries is immature, extant discussion is necessarily both qualitative and speculative. Of course, no state has ever attempted nuclear counterforce against another; this lack of data makes discussion further removed from reality.

This paper addresses these gaps by programmatically simulating a US nuclear counterforce effort assisted by deep learning against the Chinese arsenal. The focus on deep learning follows the previous paper in this dissertation, serving as a clear technical bound on how the AI works; the use of simulation extends several previous works assessing the US-China nuclear balance.[286] We operationalize this approach by focusing on

[284] Michael C. Horowitz, Paul Scharre, and Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," *arXiv*, December 13, 2019, https://arxiv.org/ftp/arxiv/papers/1912/1912.05291.pdf; Vincent Boulanin, Lora Saalman, Petr Topychkanov, Fei Su, and Moa Peldan Carlsson, "Artificial Intelligence, Strategic Stability and Nuclear Risk," *SIPRI*, June 2020, https://www.sipri.org/sites/default/files/2020-06/artificial_intelligence_strategic_stability_and_nuclear_risk.pdf, especially 23-7; Vincent Boulanin, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume I: Euro-Atlantic Perspectives*, (SIPRI, May 2019), https://fas.org/sgp/crs/natsec/R45178.pdf; Lora Saalman, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume II: East Asian Perspectives* (SIPRI: October 2019). Available online: https://www.sipri.org/publications/2019/other-publications/impact-artificial-intelligence-strategic-stability-and-nuclear-risk-volume-ii-east-asian; Petr Topychkanov, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume III: South Asian Perspectives* (SIPRI: April 2020). Available online: https://www.sipri.org/publications/2020/other-publications/impact-artificial-intelligence-strategic-stability-and-nuclear-risk-volume-iii-south-asian; Vincent Boulanin and Maaike Verbruggen, "Mapping the Development of Autonomy in Weapon Systems," *SIPRI*, 2017, https://www.sipri.org/publications/2017/other-publications/mappingdevelopment-autonomy-weapon-systems; Rafael Loss and Joseph Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?", *War on the Rocks*, September 19, 2019, https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/; Joseph Johnson, "MAD in an AI Future?", *Lawrence Livermore National Laboratory*, June 14, 2019, https://www.osti.gov/servlets/purl/1527284; Rafael Loss, "Artificial Intelligence, the Final Piece to the Counterforce Puzzle?", *Lawrence Livermore National Laboratory*, September 30, 2019, https://www.osti.gov/servlets/purl/1568008; Zachary Kallenborn, "AI Risks to Nuclear Deterrence Are Real," *War on the Rocks*, October 10, 2019, https://warontherocks.com/2019/10/ai-risks-to-nuclear-deterrence-are-real/; Edward Geist and Andrew J. Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?," *RAND*, 2018, https://www.rand.org/pubs/perspectives/PE296.html; "AI and the Military: Forever Altering Strategic Stability," *T4GS Reports*, February 13, 2019, https://www.tech4gs.org/uploads/1/1/1/5/111521085/ai_and_the_military_forever_altering_strategic_stability__t4gs_research_paper.pdf; Kenneth Payne, "Artificial Intelligence: A Revolution in Strategic Affairs?," *Survival* 60 (2018): 7-32; James Johnson, "Artificial Intelligence in Nuclear Warfare: A Perfect Storm of Instability?," *Washington Quarterly* 43.2 (2020): 197–211; Mark Fitzpatrick, "Artificial Intelligence and Nuclear Command and Control," *Survival* 61.3 (2019): 81-92; Jessica Cox and Heather Williams, "The Unavoidable Technology: How Artificial Intelligence Can Strengthen Nuclear Stability," *The Washington Quarterly*, 44.1 (2021): 69-85.

[285] For example, Horowitz et al. discuss first-wave symbolic AI in "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," Loss and Johnson discuss deep learning, a modern machine learning method, in "Will Artificial Intelligence Imperil Nuclear Deterrence?", and Payne discusses the arrival of thinking, psychologically autonomous future AI in "Artificial Intelligence: A Revolution in Strategic Affairs?".

[286] Previous works using simulation in the US-China context include Andy Halterman and Rachel Tecott, "The Case for Campaign Analysis: A Method for Studying Military Operations," *International Security* (forthcoming); Wu Riqiang, "Living with Uncertainty: Modeling China's Nuclear Survivability," *International Security* 44.4 (2020), 84-

how deep learning affects the detectability, identification, and tracking of Chinese strategic road-mobile missiles; we discuss the model in detail below.

Our paper intends several contributions. First, we weigh in on a narrow technical debate with significant stakes – the question of whether mobile nuclear platforms can be detected, identified, and tracked.[287] Second, our paper contributes to the study of artificial intelligence by political science. AI generally presents "legibility" issues – since the technology is immature and applications are many, getting a methodological handle on AI as a variable requires significant conceptual work.[288] Our operationalization in this paper does some of that work.

In addition, a significant ongoing debate about AI is whether it will diffuse power more generally, accelerating the arrival of multipolarity. On some views, the AI field's openness and commercial importance means the military benefits of the technology will be easily obtainable by other states; on other views, Chinese authoritarianism, population size, and military-civil fusion make it well-positioned to use AI to offset American power.[289] This paper provides evidence that AI's strategic uses will also be important for assessing AI's impact on the balance of power: even if both China and the United States obtain certain raw AI capabilities, analysis needs to account for whether those capabilities can be put to good use. Here, US AI capabilities threaten Chinese second-strike because of relative arsenal size and US willingness to entertain counterforce efforts, regardless of whether China also possesses equivalent deep learning capabilities. Relatedly, a small literature asks how AI will affect broader structural variables in international affairs, such as the offense-defense balance; our results support the theory that AI will promote offense-dominance by advantaging finders over hiders.[290]

---

118; and Eric Heginbotham, Michael Nixon, Forrest E. Morgan, Jacob L. Heim, Jeff Hagen, Sheng Li, Jeffrey Engstrom, Martin C. Libicki, Paul DeLuca, David A. Shlapak, David R. Frelinger, Burgess Laird, Kyle Brady, and Lyle J. Morris, "The U.S.-China Military Scorecard: Forces, Geography, and the Evolving Balance of Power 1996-2017," *RAND*, 2015, , 285-417. Work by Lieber and Press here is seminal, although they do not specifically model the US-China balance. See Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security* 41.4 (2017): 9-49; ---, "The End of MAD? The Nuclear Dimension of U.S. Primacy," *International Security* 30.4 (2006): 7-44.

[287] For convenience, we use "detectability" in this paper as shorthand for both finding and tracking road-mobile missiles, although these are distinct intelligence tasks in practice (and are modeled separately in our simulation).

[288] Michael C. Horowitz, "Do Emerging Military Technologies Matter for International Politics?", *Annual Review of Political Science* 23 (2020), 385-400.

[289] Robert O. Work and Greg Grant, "Beating the Americans at Their Own Game: An Offset Strategy with Chinese Characteristics," *CNAS*, June 6, 2019, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Report-Work-Offset-final-B.pdf?; Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1.3 (2018), https://tnsr.org/2018/05/artificial-intelligence-international-competition-and-the-balance-of-power/; Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt Publishing Company, 2018); Remco Zwetsloot, Helen Toner, and Jeffrey Ding, "Beyond the AI Arms Race: America, China, and the Dangers of Zero-Sum Thinking," *Foreign Affairs*, November 16, 2018, https://www.foreignaffairs.com/reviews/review-essay/2018-11-16/beyond-ai-arms-race; James Johnson, "The end of military-techno Pax Americana? Washington's strategic responses to Chinese AI-enabled military technology," *The Pacific Review* (2019).

[290] Horowitz et al., "A Stable Nuclear Future?"; Payne, "Artificial Intelligence"; Remco Zwetsloot and Allan Dafoe, "Thinking About Risks From AI: Accidents, Misuse and Structure," *Lawfare*, February 11, 2019, https://www.lawfareblog.com/thinking-about-risks-ai-accidents-misuse-and-structure.

Finally, we push the envelope on the sophistication of simulations used in campaign analyses. Previous models primarily use Monte Carlo sampling to allow specified quantities to vary according to some set distribution. While useful, these models also naturally limit themselves to spreadsheet-style analysis. In contrast, we implement a quasi-geographic model in Python which operationalizes each TEL as a state machine, enabling modeling of the US intelligence process over many time-steps. This not only allows us to sample from a richer array of possible simulated world-states, but additionally permits capturing complicated, emergent, over-time effects from the progressive interaction of various modeled objects.

The remainder of the paper proceeds as follows: first, we unpack the debate over whether, why, and how AI may impact the viability of first-strike counterforce. Second, we describe the programmatic model we use to test these mechanisms, and then present the model's results. Finally, we discuss the broader implications of our model's results, including for what countermeasures states may adopt to stave off the impact of AI on their arsenals, and briefly conclude.

## Artificial Intelligence and Nuclear Counterforce

Scholars have proposed many ways in which AI could affect the possibility of nuclear counterforce. On the one hand, AI could enable fast, high-endurance drones capable of penetrating deep into enemy territory, solve "needle in a haystack" problems associated with tracking mobile platforms both on land and undersea, and increase the speed and reliability of conventional counterforce attacks, among other effects. Thus, AI could decrease survivability.

On the other hand, AI could improve early warning systems, increase the robustness of nuclear command and control systems, and enable the use of highly survivable autonomous delivery vehicles able to remain meaningfully functional past the loss of human leaders, among other effects. Thus, AI could enhance survivability.[291]

The list of theorized effects is very long, as is the list of applications enabled by AI which are said to responsible for producing such effects. The key difficulty, of course, is that there is likely some truth to the diagnosis of such effects on both sides, AI being a general-purpose technology which "goes in everything." Fortunately, we claim, most hypothesized effects in political science should be considered not to be real until proven beyond conceptual gesture; in the grand interplay of technology, after all, it is axiomatic that most effects "wash out," easily counterbalanced by some equally incidental countermeasure and thereby never of lasting interest to political scientists – to produce effects on international affairs which are notable, we ought to have some real evidence to believe some real effect exists. Consequently, we limit ourselves to examining the use of deep learning techniques to assist with intelligence-processing in the detectability, identification, and tracking of Chinese strategic road-mobile missiles.

### (1) Deep Learning and Intelligence-Processing

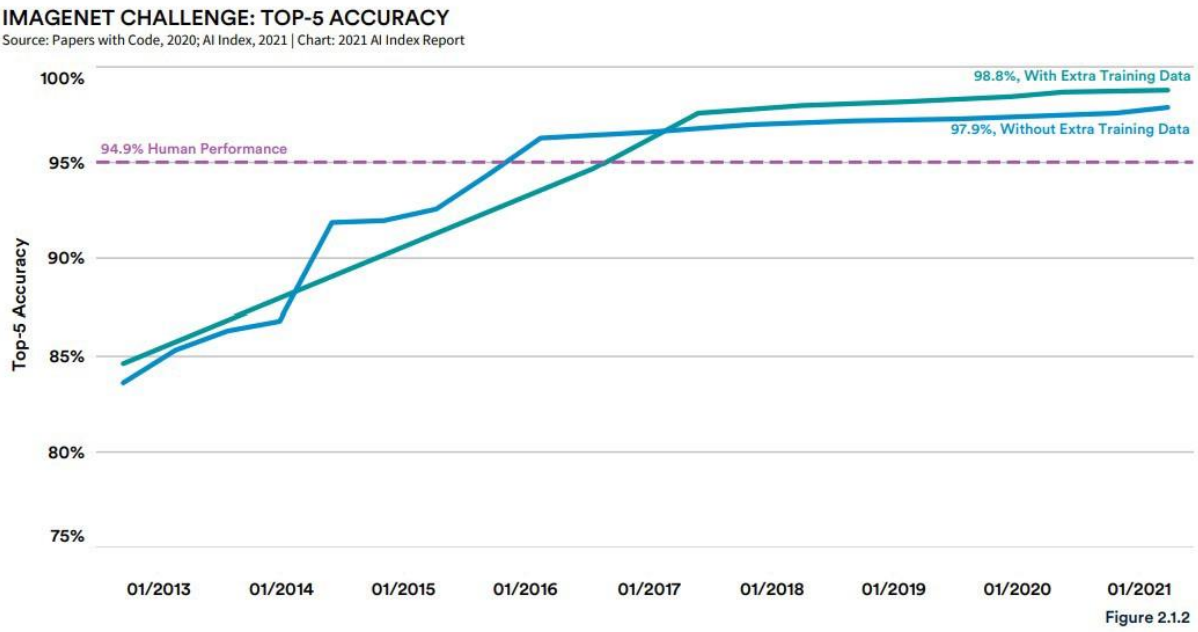
This focus has several rationales. First, deep learning for pattern recognition in colossal-scale data streams is a proven application of AI which has already been demonstrated to be technically possible. Consequently, there is no need to speculate about, for example, theoretical delegation to future emotionless AI decision-

[291] Horowitz, Scharre, and Velez-Green, "A Stable Nuclear Future?"; Boulanin et al., "Artificial Intelligence, Strategic Stability and Nuclear Risk"; ---, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume I*; Saalman, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume II*; Topychkanov, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume III*; Loss and Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?"; Johnson, "MAD in an AI Future?"; Loss, "Artificial Intelligence, the Final Piece to the Counterforce Puzzle?"; Geist and Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?," *RAND*, 2018; "AI and the Military"; Payne, "Artificial Intelligence: A Revolution in Strategic Affairs?".

makers, as some commentators do. Second, although descriptively innocuous, deep learning is actually a breakthrough technologically in the ability of militaries to process intelligence at speed and at scale. As our model results below demonstrate, the application of deep learning can increase the speed at which a military locates a target of interest in already-possessed intelligence by between one and three orders of magnitude.

Critically, although efforts to use other technologies commonly referred to as AI have been applied to intelligence processing before, none have reached near-human (or, in the case of deep learning, often super-human) levels of intelligence, due to the difficulty of generating a mathematical formalizing which captures what humans do when they, for example, identify a cat as a cat, or a SAM site as such. Due to widespread application of deep learning to image, text, video, and other modalities beginning in the 2010s, however, deep learning classifiers now exceed human performance on many standard benchmarks of competence at these tasks.

**ImageNet Challenge Progress**[292]




Consequently, at the level of military tasks, substitution of deep learning algorithms for human intelligence analysts in many modalities has become increasingly feasible; AI can be tasked with locating targets of interest, noticing discrepancies in behavior, or combing through endless quantities of online text to look for terrorist intentions.[293]

---

[292] Daniel Zhang, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, James Manyika, Juan Carlos Niebles, Michael Sellitto, Yoav Shoham, Jack Clark, and Raymond Perrault, "Artificial Intelligence Index Report 2021," *Stanford University Center for Human-Centered Artificial Intelligence*, 2021, https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report_Master.pdf.

[293] Innovation in the algorithmic architectures underlying deep learning's abilities in these domains continue. For imagery, the chief innovation for deep learning's classifying abilities was convolutional neural nets (CNNs), which encoded a prior that pixels near each other tended to contain relevant information in images; for text, the key architectural improvement has been the Transformer, which also underly key recent breakthroughs in the quality of machine translation in BERT, and text generation in GPT-2 and GPT-3. For the seminal papers involved, see Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural

Third, though mundane-sounding, enhanced speed and scale in intelligence-processing strike at a key bottleneck in the US counterforce process vis-à-vis China: that of detectability. Surprisingly, detectability appears to decisively inform analysts' views on whether China's nuclear arsenal is survivable, but has not previously been modeled. Instead, detectability or lack thereof is a stipulated assumption in extant models.

Reviewing briefly, Keir Lieber and Daryl Press (2006) assess that the US maintains a first-strike capability against both Russia and China by assuming, without explicitly modeling the question, that detectability is trivial.[294] In contrast, Eric Heginbotham et al. (2015) assess that the Chinese arsenal has become increasingly survivable because they assume, also without modeling, that Chinese mobile assets are invulnerable.[295]

Most recently, Wu Riqiang (2020) uses point estimates for the detectability of different Chinese fixed assets to model US nuclear counterforce over a 25-year period, but does not cite sources for these estimates.[296] Rachel Tecott and Andy Halterman (2021) correctly note the arbitrary nature of Wu's detectability point estimates, and perform a sensitivity analysis by sampling from reasonable ranges – for example, per iteration, they draw the probability a launch site is detected from the range (10%, 90%).[297] In these analyses, however, one's assumed degree of detectability is still decisive – drawing a higher probability makes China vulnerable to a US first-strike, while low detectability implies the inverse.

In short, if we could narrow the range for how detectable we believe Chinese mobile nuclear assets are, this would dominate our conclusion about Chinese survivability in one direction or another. Intuitively, at a very simple level, this is actually rather unsurprising – the US nuclear arsenal is large relative to China's, so as long as any US counterforce effort can trade its weapons for China's at some sane exchange rate, counterforce will look possible in these models. As the Chinese academic Li Bin once somewhat scathingly commented about Lieber and Press (2006),

> Basic arithmetic alone will certify that thousands of nuclear missiles should be able to destroy a couple dozen [missiles] … [but] the calculations in the paper are based on a fundamentally unrealistic assumption: that is, [that] the United States can detect and locate all Russian and Chinese long-range nuclear weapons.[298]

Our paper addresses this gap in the literature by modeling, from first principles, whether the United States can detect China's road-mobile missiles using its various intelligence assets. We assess this question both with and without AI. We focus on road-mobile missiles because both Chinese and American analysts

---

Networks," Proceedings of the 25th International Conference on Neural Information Processing Systems 1, December 2012, 1097–105. Available online: https://dl.acm.org/doi/10.5555/2999134.2999257; and Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention is All You Need," arXiv, December 16, 2017, https://arxiv.org/pdf/1706.03762.pdf. For an overview of deep learning applied to SAR data specifically, see C. P. Schwegmann, W. Kleynhans, and B. P. Salmon, "The development of deep learning in synthetic aperture radar imagery," paper presented at the International Workshop on Remote Sensing with Intelligent Processing (RSIP), 2017, China, https://doi.org/10.1109/RSIP.2017.7958802.

[294] Lieber and Press, "The End of MAD?".

[295] Heginbotham et al., "The U.S.-China Military Scorecard."

[296] Wu, "Living with Uncertainty," 92-3.

[297] Tecott and Halterman, "The Case for Campaign Analysis," 25-6.

[298] Li Bin, "Paper Tiger with Whitened Teeth," *China Security* (2006), 78-89. Available online: https://www.issuelab.org/resources/437/437.pdf.

mostly agree China's fledgling SSBN capability is unlikely to become survivable in the medium-term; the undersea balance strongly favors the United States.[299]

### (2) Intelligence-Processing and Counterforce

So, how does AI help with detectability? The core of the problem is as follows: US collection of data has far outpaced the ability to digest it.[300] As US sensor platforms have proliferated, so has the demand for skilled human labor to process the resulting intelligence. Intelligence analysts reportedly can require a full day simply to exploit 6 to 12 percent of imagery data for one city; one defense official has projected that the intelligence community would require 8 million analysts to manually exploit its imagery data alone in 2037. Consequently, the United States is actively pursuing AI-empowered intelligence-processing capabilities to find and track mobile nuclear assets, both on land and at sea.[301] In 2018, Reuters quoted US officials noting that "there are multiple classified programs now under way to explore how to develop AI-driven systems ... scouring huge amounts of data, including satellite imagery, with a speed and accuracy beyond the capability of humans."[302] Since computational power is easily purchased much more cheaply than human analysts, highly accurate classifiers would enable the US intelligence community to run any number of iterations of its code over various intelligence modalities. Although this seems to us well within the capability of existing deep learning systems, we pause to consider two objections:

#### (a) What about data quality?

Rafael Loss and Joseph Johnson argue in several fora that data quality issues will prevent AI from usefully classifying Chinese road-mobile transporter-erector launchers (TELs), but we disagree.[303]

---

[299] Tong Zhao, "Tides of Change: China's Nuclear Ballistic Missile Submarines and Strategic Stability," *Carnegie Endowment for International Peace*, 2018, https://carnegieendowment.org/files/Zhao_SSBN_final.pdf, 26-7; Wu Riqiang, "Survivability of China's Sea-Based Nuclear Forces," *Science and Global Security* 19.2 (2011), Charles L. Glaser and Steve Fetter, "Should the United States Reject MAD? Damage Limitation and U.S. Nuclear Strategy toward China," *International Security* 41. 1 (2016), 70-2; Owen R. Cote Jr., "Assessing the Undersea Balance Between the U.S. and China," *SSP Working Paper*, 2011, https://www.usni.org/sites/default/files/inline-files/Undersea%20Balance%20WP11-1.pdf; see also ---, "The Third Battle: Innovation in the U.S. Navy's Silent Cold War Struggle with Soviet Submarines," *Naval War College Newport Papers* 16 (2003), https://digital-commons.usnwc.edu/newport-papers/38/, and Austin Long and Brendan Rittenhouse Green, "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategic Studies* 38 (2015): 38-73. We leave modeling the undersea balance to future work.

[300] Mark Pomerleau, "Can the intel and defense community conquer data overload?," C4ISRNET, September 5, 2018, https://www.c4isrnet.com/intel-geoint/2018/09/05/can-the-intel-and-defense-community-conquer-data-overload/?.

[301] Marcus Weisgerber, "The Increasingly Automated Hunt for Mobile Missile Launchers," *Defense One*, April 26, 2016, https://www.defenseone.com/technology/2016/04/increasingly-automated-hunt-mobile-missile-launchers/127864/; Patrick Tucker, "How AI Will Transform Anti-Submarine Warfare," *Defense One*, July 1, 2019, https://www.defenseone.com/technology/2019/07/how-ai-will-transform-anti-submarine-warfare/158121/; Jeremy Hsu, "Wanted: AI That Can Spy," *IEEE Spectrum*, November 2017, https://spectrum.ieee.org/aerospace/satellites/wanted-ai-that-can-spy;

[302] Phil Stewart, "Deep in the Pentagon, a secret AI program to find hidden nuclear missiles," *Reuters*, June 5, 2018, https://www.reuters.com/article/us-usa-pentagon-missiles-ai-insight/deep-in-the-pentagon-a-secret-ai-program-to-find-hidden-nuclear-missiles-idUSKCN1J114J.

[303] Loss and Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?"; Johnson, "MAD in an AI Future?"; Loss, "Artificial Intelligence, the Final Piece to the Counterforce Puzzle?".

First, the data imbalance failure modes they mention are trivially easy to avoid. The US likely has a reasonable quantity of images of TELs, which have been paraded through China and photographed on the Internet, but even were China to deploy a novel TEL type during a crisis, one can simply add weights to datasets so that false negatives are costlier than false positives. Further, it would not be difficult to retrain one's model on a daily basis as new intelligence arrived of the new TEL. Alternatively, one could synthetically augment the number of positive images in any number of ways; rotation, shifting, and cropping, for example, would not increase the number of false positives, since they would represent potentially real pictures of TELs heading in different directions.[304]

Second, they argue AI cannot understand vehicle function from form, but this is a non sequitur. For intelligence processing, the AI merely needs to sort images; it does not need to have any understanding of how missile launchers work. If AI can never locate TELs because they are indistinguishable from trucks, this would also be equally true of human analysts.

Third, they argue AI cannot deal with the curse of dimensionality. This is a strange criticism, since deep learning's significant advances over the state-of-the-art have resulted precisely from its unique ability to deal with the curse of dimensionality.[305] Further, the curse of dimensionality simply does not apply to the examples they cite. The curse of dimensionality is that phenomenon in statistics that as dimensionality increases, the difficulty of estimating a function grows exponentially due to combinatorial explosion.[306] Increasing picture resolution, however, is not equivalent to adding points in additional dimensions. One does not store 100x100 pixel images as 10,000-dimensional vectors. Contextually, deep convolutional neural networks (CNNs) exactly work by introducing spatial dependency between nearby pixels. In the lab, CNN-based classifiers run over SAR data for automatic target recognition (ATR) have already achieved accuracy in excess of 95%.[307]

Finally, they argue that AI will be unable to discard unimportant features. However, deep CNNs *do* in fact possess this ability – that is, in fact, the reason to use them. The oft-repeated buzzword about deep learning is that it makes feature engineering, or the careful selection of what features to feed into the model, unnecessary. This criticism of theirs is thus especially strange.

**(b) What about the "Scud hunt"?**

During the oft-referenced "Scud hunt," the United States launched around 2,500 sorties intended to hunt Scud launchers, but failed to confirm kills against any of the 30 Iraqi TELs, despite 88 Scud missiles being

[304] Ryan J. Soldin, Douglas N. MacDonald, Matthew Reisman, Latisha R. Konz, Roger Rouse, and Timothy L. Overman, "HySARNet: a hybrid machine learning approach to synthetic aperture radar automatic target recognition," *paper presented at SPIE Defense + Commercial Sensing*, 2019, United States, https://doi.org/10.1117/12.2518155

[305] For the explanation in the standard introductory textbook, see Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (MA: MIT Press, 2016), 157-161. See also an excellent intuitive explanation at Chris Olah, "Neural Networks, Manifolds, and Topology," *colah's blog*, April 6, 2014, http://colah.github.io/posts/2014-03-NN-Manifolds-Topology/.

[306] Lei Chen, "Curse of Dimensionality," in *Encyclopedia of Database Systems*, ed. Ling Liu and M. Tamer Ozsu (Boston: Springer, 2009).

[307] Soldin et al., "HySARNet"; Andrew Profeta, Andres Rodriguez, and H. Scott Clouse, "Convolutional neural networks for synthetic aperture radar classification," *paper presented at SPIE Defense + Security*, 2016, United States, https://doi.org/10.1117/12.2225934.

launched.[308] The Scud hunt is thus often cited to illustrate the difficulty of counterforce, including in discussions about AI.

However, compared to modern US counterforce against China, the Scud hunt is disanalogous even before adding AI. First, the Scud TEL had a much smaller intelligence signature – it was ten times lighter and four times smaller than the DF-31A, and was not accompanied by various unique support vehicles.[309] Chinese TELs include service trucks purposed for fire control, power, power distribution, aiming, and inspection.[310] The presence of these trucks themselves – or their distinct signatures – could all be features that AI algorithms could learn to detect.

Second, with US counterforce against China, less precision is required of any intelligence signature – US nuclear weapons would destroy a much larger area than the conventional munitions used against Scud TELs.[311] This enables our model's use of barrage attacks against Chinese TELs, which tolerate uncertainty or targeting delay simply by destroying a wide area. In contrast, although Scud TELs were visually sighted 42 times during the Scud hunt, strike aircraft were unable all but 8 times to find the TELs again after arriving in the area.[312]

Third, the United States and its coalition partners failed to undertake extensive "intelligence preparation of the battlefield" (IPB); analysts were unsure even how many TELs existed, let alone where they were based, let alone where they would likely hide based on previous behavior. In contrast, the US devotes significant effort to monitoring Chinese TELs, as we discuss further below.[313] Fourth, the sensor modalities used during the Scud hunt were qualitatively much worse than those available today – search modalities consisted entirely of aircraft and special operations forces (SOF) operating on foot, with both sometimes defaulting to using own literal eyesight to spot TELs.[314]

Adding AI, our model illustrates how AI-assisted intelligence-processing shores up exactly those deficiencies which produced the failed Scud hunt. First, during the Scud hunt, the typical delay between target identification and ordnance delivery exceeded 50 minutes.[315] Consequently, cued aircraft were often unable to find their targets. In contrast, our simulation results graphed below show exactly that AI's key effect is to accelerate intelligence-processing, enabling the US to attack a much smaller area to defeat a given TEL. Second, AI enables broad-area search modalities to be useful for mobile targets. Using aircraft

[308] About 1,000 of these failed to locate TELs, and so instead attacked other targets of opportunity. See Barry D. Watts and Thomas A. Keaney, "Effects and Effectiveness," in *Gulf War Air Power Survey, Volume II: Operations and Effects and Effectiveness*, ed. Eliot A. Cohen (Washington, DC: US Government Publishing Office, 1993), 330-40. Available online: https://apps.dtic.mil/dtic/tr/fulltext/u2/a279742.pdf.

[309] Glaser and Fetter, "Should the United States Reject MAD?," 67.

[310] Wu, "Certainty of Uncertainty," 586.

[311] Ibid.

[312] William Rosenau, "Special Operations Forces and Elusive Enemy Ground Targets: Lessons from Vietnam and the Persian Gulf War," *RAND*, 2001, https://www.rand.org/pubs/monograph_reports/MR1408.html, 34.

[313] Ibid., 32; Wu Riqiang, "Certainty of Uncertainty: Nuclear Strategy with Chinese Characteristics," *Journal of Strategic Studies* 36.4 (2013), 488, footnote 38.

[314] Rosenau, "Special Operations Forces and Elusive Enemy Ground Targets," 38, 42.

[315] Ibid., 39.

and SOF to find Scuds was like – and in many cases, exactly was – walking through a desert with a flashlight looking for small objects. In contrast, using AI to sort through vast quantities of intelligence before passing on cues to human analysts would be like knowing, after a few minutes, the three possible locations those small objects could be.

## Modeling US Nuclear Counterforce Against China

Having described how AI could assist with US counterforce, we now turn to describing the model within which that assistance occurs. At a high level, our code simulates the kill chain from end to end for US nuclear counterforce efforts against the Chinese arsenal, separately modeling each of "find, fix, and finish." We then evaluate any remaining Chinese weapons against US missile defenses; this enables us to output whether China did, in fact, successfully retaliate against any given counterforce attempt.[316]

Our baseline case implements AI-assisted "finding" and "fixing" against low, medium, and high alert levels; these Chinese alert levels, which we describe below, control TEL behavior, including the frequency and roaming behavior of deterrent patrols, the strictness of emissions control, what percentage of Chinese weapons are mated, and the density of Chinese decoys, among other details. We simulate both broad-area and cued collection modalities for the United States. For the former, this includes imagery satellites (EO), radar satellites (SAR/GMTI), and patrolling standoff aircraft operating near China's coast (P-8s, EP-3s); for the latter, this includes unattended ground sensors, hyperspectral imagers, SIGINT intercepts, and stealthy penetrating UAVs (RQ-180s). To more realistically model these collection modalities, we also simulate various stochastic environmental features which likely affect US intelligence efforts. Our simulation includes a day/night cycle, cloud cover varying over hours, and geographically specific estimates of Chinese truck density as an input into false-positive TEL sightings.

Our simulation of US "finding" and "fixing" outputs target sets for US counterforce efforts, including both fixed and mobile targets; our code then evaluates the sufficiency of the US arsenal at each time-step. We launch US SLBMs and ICBMs in barrage attacks against Chinese TELs, then launch any surviving Chinese weapons against CONUS, assessing whether they penetrate US missile defenses. This enables us to graph, over the course of the simulation, windows of opportunity where US counterforce efforts would succeed, if any. Below, we present results from twelve runs which simulate a 24-hour period minute by minute, but no programmatic limitation would prevent longer simulations. Our code is also highly configurable across all of these parameters, enabling easy exploration of excursion cases; below, we evaluate the implications of the ongoing miniaturization of satellite technology, as well as of Chinese expansion of its nuclear arsenal. We also run the model with and without AI, allowing us to illustrate its impact.

The remainder of this section discusses the simulation's parameters for the Chinese nuclear arsenal, TEL behavior under different alert levels, US intelligence collection, US intelligence processing, US nuclear counterforce attacks, and Chinese retaliation.

### (1) Chinese Nuclear Forces

In 2019, according to the oft-cited *Nuclear Notebook*, China possessed around 290 nuclear warheads, intended for delivery via land-based ballistic missiles, submarine-launched ballistic missiles (SLBMs), and aircraft. However, many of these warheads were intended for regional use. Approximately 138 warheads were attached to delivery systems which range US soil, including territories and peripheral states; of these, around 80 were able to reach the continental United States (CONUS) proper from their deployment areas.

[316] While we focus the most effort on Chinese road-mobile missiles, we account for other parts of the Chinese arsenal through parameter adjustment.

Roughly 48 of those 80 were road-mobile missiles.[317] As of November 2020, according to the Nuclear Threat Initiative (NTI), China has approximately 64 such missiles.[318] We combine data sources to produce the below table of Chinese TELs and associated bases.[319]

**Chinese Nuclear-Capable TELs, 2019**

| name | latitude | longitude | missile_count | DF_31A | DF_31AG | DF_26[320] | DF_31 | DF_21AE |
|---|---|---|---|---|---|---|---|---|
| Datong | 36.95 | 101.67 | 6 | 0 | 0 | 0 | 6 | 0 |
| Chizhou | 30.69 | 117.9 | 12 | 0 | 0 | 0 | 0 | 12 |
| Dalian | 39.3 | 122.07 | 12 | 0 | 0 | 0 | 0 | 12 |
| Leping | 28.98 | 117.12 | 12 | 0 | 0 | 0 | 0 | 12 |
| Nanyang | 33.01 | 112.41 | 12 | 12 | 0 | 0 | 0 | 0 |
| Shaoyang | 27.25 | 111.39 | 12 | 0 | 12 | 0 | 0 | 0 |
| Tianshui | 34.53 | 105.91 | 12 | 0 | 12 | 0 | 0 | 0 |
| Yuxi | 24.36 | 102.49 | 24 | 24 | 0 | 0 | 0 | 0 |
| Hancheng | 35.48 | 110.45 | 6 | 2 | 2 | 0 | 2 | 0 |
| Korla | 41.69 | 86.17 | 36 | 0 | 0 | 36 | 0 | 0 |
| Jianshui | 23.73 | 102.87 | 36 | 0 | 0 | 36 | 0 | 0 |
| Xinyang | 32.17 | 114.13 | 36 | 0 | 0 | 36 | 0 | 0 |
| Qingyuan | 23.68 | 113.18 | 36 | 0 | 0 | 36 | 0 | 0 |

**Mobile Missiles Ranging CONUS**

China's aforementioned 64 strategic mobile missiles ranging CONUS are single-shot transporter-erector-launchers (TELs), represented entirely by the DF-31A (CSS-10 Mod 2) and DF-31AG (CSS-10 Mod 3).[321] In 2019, China had approximately 24 launchers each of both types; they carry the same missile ranging 11,200 km.[322] As of November 2020, NTI reports China has approximately 38 DF-31As and 26 DF-31AGs.[323]

[317] Hans M. Kristensen and Matt Korda, "Chinese nuclear forces, 2019," *Bulletin of the Atomic Scientists* 75.4 (2019): 171-8.

[318] Jeffrey Lewis, David Joel La Boon, and Decker Eveleth, "China's Growing Missile Arsenal and the Risk of a 'Taiwan Missile Crisis,'" *Nuclear Threat Initiative*, November 18, 2020, https://www.nti.org/analysis/articles/chinas-growing-missile-arsenal-and-the-risk-of-a-taiwan-missile-crisis/.

[319] As noted above, we use these estimates of current Chinese inventory in our baseline case, but explore Chinese nuclear arsenal expansion in excursion cases.

[320] Dual-capable.

[321] The DF-41 would also fall into this category, but is not yet operational.

[322] The DF-31AG is simply an improved launcher. Kristensen and Korda, "Chinese nuclear forces, 2019," 172-5.

[323] Lewis et al., "China's Growing Missile Arsenal and the Risk of a 'Taiwan Missile Crisis.'"

**Mobile Missiles Ranging US Soil**
China additionally possesses approximately 46 mobile missiles which range US territories, Hawaii, or Alaska, but not CONUS. These missiles comprise:

> ● **40 DF-26s (no NATO designation).** China has approximately 200 DF-26 launchers, capable of carrying both conventional and nuclear payloads.[324] Kristensen and Korda estimate approximately 40 would arm themselves with nuclear warheads in a crisis, with a range of 4,000 km; distinguishing them on the road may be difficult. This may require counterforce efforts to destroy all DF-26 launchers, if distinguishing intelligence streams are unavailable. Unlike China's DF-21 variants, where the nuclear and non-nuclear launchers have differences easily observable in satellite imagery, the DF-26 is simply one dual-use launcher.[325] These could reach Guam, but not CONUS.[326]
>
> ● **6 DF-31s (CSS-10 Mod 2).** China has about 6 DF-31 launchers with a range of about 7,200 km. They are not deployed in areas which would enable them to reach CONUS, but could target Guam.[327] Theoretically, DF-31s deployed in Heilongjiang, China's northeast corner, could just reach parts of the American west coast.

We focus on the combined target set of missiles ranging either CONUS or US soil in our baseline case, but US planners could theoretically set the threshold for successful counterforce at eliminating all nuclear weapons able to target even US overseas bases and/or regional allies.

**Mobile Missiles Ranging US Allies**
Finally, China also possesses approximately 40 missiles total for the DF-21A (CSS-5 Mod 2) and DF-21E (CSS-5 Mod 6). These are the nuclear variants of the dual-capable DF-21 family, and have a range of 2,150 km.[328] As above, the variants are distinguishable. They can reach neither Guam nor CONUS, but could threaten US regional allies.

## (2) Chinese TEL Behavior

How do Chinese transporter-erector-launchers (TELs) behave under different alert levels? Unfortunately, with some notable exceptions, scarce explicit information is available.[329] Consequently, we follow Heginbotham et al. (2015) and Wu (2020) in conceptualizing several alert levels appropriate to our simulation based on reasonable assumptions about Chinese strategy.

Fundamentally, China faces tradeoffs between TEL survivability and other goals. Trivially, were China only seeking to maximize TEL survivability, it could purchase an enormous number. In practice, potential

[324] Most of the DF-26 launchers are purely conventional.

[325] James Acton, "The Evolution of Ambiguous Weapons," *Carnegie Endowment for International Peace*, April 9, 2020, https://carnegieendowment.org/2020/04/09/evolution-of-ambiguous-weapons-pub-81449.

[326] Kristensen and Korda, "Chinese nuclear forces, 2019," 172-5. "DF-26 (Dong Feng-26)," *CSIS Missile Defense Project*, June 23, 2020, https://missilethreat.csis.org/missile/dong-feng-26-df-26/.

[327] Kristensen and Korda, "Chinese nuclear forces, 2019," 172-5.

[328] Kristensen and Korda, "Chinese nuclear forces, 2019," 172; "DF-21 (Dong Feng-21 / CSS-5)," *CSIS Missile Defense Project*, January 2, 2020, https://missilethreat.csis.org/missile/df-21/.

[329] Heginbotham et al., "The U.S.-China Military Scorecard," 288-9.

defensive practices like emissions control (EMCON), TEL hardening or stealth covers, or off-road capability may also pose difficulties for maintaining nuclear command and control, run against PLA culture, and/or simply cost money presently allocated elsewhere. For historical reasons dating back to Mao, for example, the CCP strongly prioritizes negative over positive control – in relative terms, it prefers the "never" branch of the always/never dilemma.[330] Chinese elites are likely to strongly resist pre-delegation to field commanders, even when tactically optimal.[331] Strategically, China's declared "no first use" policy looks more credible the less TELs are ready to fire.[332]

Consequently, China presently keeps its missiles de-mated and its TELs in garrison during peacetime, only dispersing launchers during crises as deliberate nuclear signaling.[333] Relative to continuous deterrent patrols, this obviously decreases TEL survivability, but serves as a costly and thus arguably credible signal of Chinese nuclear intentions. Under this peacetime level of "day-to-day alert," US first-strike counterforce would not require finding and tracking dispersed TELs, but merely striking various possible combinations of fixed assets, such as the TEL-associated bases themselves, the warhead storage base or its associated rail lines, prepared forward and launch sites for default TEL deployment, and so on.[334] Since this alert level involves no application of AI against mobile systems, we model it only as a beginning state for shifting to other alert levels. That is, even our "low alert" scenario assumes higher readiness of nuclear forces than China actually practices, a China-favoring assumption already giving credit for some level of adaptation.[335]

Instead, given the above, we conceptualize Chinese alert levels as a progressively increasing willingness to prioritize TEL survivability over other goods. Applying this idea to various available levers allows us to formulate corresponding TEL behaviors. We thus model three alert levels: low, medium, and high.

**TEL Behavior by Alert Level**

| | Low | Medium | High |
|---|---|---|---|
| **Speed (km/h)** | 20 | 40 | 69 |
| **Mating %** | 16.67% | 50% | 100% |
| **Roam %** | 25% | 75% | 100% |
| **EMCON %** | 0% | 50% | 100% |

[330] M. Taylor Fravel, *Active Defense: China's Military Strategy Since 1949* (NJ: Princeton University Press, 2019), 236-70; Fiona S. Cunningham and M. Taylor Fravel, "Assuring Assured Retaliation: China's Nuclear Posture and U.S.-China Strategic Stability," *International Security* 40.2 (2015), 39.

[331] M. Taylor Fravel and Evan S. Medeiros, "China's Search for Assured Retaliation: The Evolution of Chinese Nuclear Strategy and Force Structure," *International Security* 35.2 (2010), 74-5.

[332] Wu, "Certainty of Uncertainty," 590.

[333] Mark A. Stokes, "China's Nuclear Warheads Storage and Handling System," *Project 2049 Institute*, March 12, 2010, https://project2049.net/wp-content/uploads/2018/05/chinas_nuclear_warhead_storage_and_handling_system.pdf; Wu, "Living with Uncertainty"; ---, "Certainty of Uncertainty," 586-7; Li Bin, "Tracking Chinese Strategic Mobile Missiles," *Science and Global Security* 15.1 (2007), 7-11.

[334] Wu, "Living with Uncertainty."

[335] Stokes, "China's Nuclear Warhead Storage and Handling System"; see also discussion in Glaser and Fetter, "Should the United States Reject MAD?," 64-5.

| | | | |
|---|---|---|---|
| **Endurance (h/day)** | 8 | 16 | 24 |
| **Endurance (km)** | 500 | 750 | 1000 |
| **Decoy:Real** | 0:1 | 1:1 | 2:1 |
| **Base Tethering?** | Y | Y | N |

**Speed**

How fast do the TELs drive? Extant scholarly works report significantly different likely Chinese TEL speeds, ranging from 20 km/h to 90 k/h.[336] Since the reported speed of the DF-31 TEL vehicle is 69 km/h, we use that as a maximum.[337] Although 20 km/h seems intuitively slow, this sometimes corresponds to Chinese road regulations for TEL-sized trucks; additionally, TEL accidents are likely seen as particularly costly, especially if they are carrying nuclear weapons.[338] Consequently, we take the 20 km/h estimate from Li (2007) at face value, but assign it to the low alert case; we assign 40 km/h to medium alert, and 69 km/h to high alert.

**Mating %**

What percentage of TELs are mated? As above, Chinese doctrine keeps nuclear warheads de-mated from TELs in peacetime.[339] For low alert, we follow Wu (2020) in assuming only 1/6$^{th}$ of warheads will be readily available, having been pre-dispersed from central storage to TEL bases. To prevent the US from simply striking the central warhead and TEL bases and easily obtaining victory, we operationalize low alert by assuming China takes the minimum precaution of mating those pre-dispersed warheads. Consequently, 16.67% of China's TELs are mated in low alert; we stipulate China mates half of TELs in medium alert, and all TELs in high alert.

**Roam %**

What percentage of TELs are roaming, as opposed to staying in base? For low and medium alert, we follow Heginbotham et al. (2015) in assigning 25% and 75% of TELs to roam, respectively.[340] Under high alert, we set 100% of TELs to roam.

**EMCON %**

What percentage of the time do TELs practice EMCON? Under low alert, drivers and staff freely communicate and receive orders and other chatter; under medium alert, we cut this by half. Under high alert, we assume pre-delegation of necessary authorities, allowing 100% EMCON.

---

[336] Li, "Tracking Chinese Strategic Mobile Missiles," 8-10; Wu, "Living with Uncertainty," Appendix, 4; Glaser and Fetter, "Correspondence," 205-6.

[337] Andrius Genys, "Hanyang HY4330," *Military Today*, January 8, 2020, http://www.military-today.com/trucks/hanyang_hy4330.htm.

[338] Li, "Tracking Chinese Strategic Mobile Missiles"; Glaser and Fetter, "Correspondence," 205.

[339] Stokes, "China's Nuclear Warheads Storage and Handling System."

[340] That is, our settings here correspond to "Posture A" and "Posture B" in the Scorecard. See Heginbotham et al., "The U.S.-China Military Scorecard," 289.

**Endurance**
How long or far can TELs drive before crews need to rest, vehicles need maintenance, or fuel tanks are empty?[341] We conceptualize longer endurance times and distances as a willingness to push crews, defer non-emergency maintenance, and carry additional fuel on-board or as part of the vehicle convoy. We assign 8, 16, and 24-hour shifts to low, medium, and high alert, respectively, with drive distances of 500, 750, and 1000 km. Under high alert, we assume refueling on the road takes half an hour.

**Decoys**
We stipulate that under low alert, China deploys no decoys; under medium alert, China deploys decoys at a 1:1 ratio. These decoys are indistinguishable from real TELs with 0.5 probability; we conceptualize them as decoys seen previously by US intelligence. Under high alert, we stipulate China deploys decoys at a 2:1 ratio, with half of deployed decoys nearly indistinguishable from real TELs (0.9).

**Base Tethering**
Under low and medium alert, we assume that TELs eventually return to their base areas, and do not arbitrarily wander the entirety of China. Under high alert, we assume that China has pre-positioned replacement crew members, fuel, food, and other maintenance supplies, thus allowing arbitrary TEL roaming without generating large, visible supply tail. In this "free roaming" mode, TELs rest for four hours every 24 hours underneath thick structures (e.g., overpasses, warehouses, tunnels) which block line-of-sight collection modalities.[342]

## (3) US Intelligence: Preparation of the Battlefield

Very substantial US peacetime intelligence efforts are devoted to collecting on others' nuclear arsenals.[343] These include broad-area remote sensing modalities, but also costlier, slower, targeted efforts, such as cultivating human sources with special knowledge of China's nuclear weapons, monitoring communications over long stretches of time, and directing satellites to look at particular areas of interest.

**Bases**
Consequently, the United States has likely already identified China's TEL-containing garrisons and pre-prepared launch sites.[344] (Minimally, it knows at least as much as we know from the open-source.) This means crises start with the US possessing knowledge of how many TELs China has of each type and which bases they are stationed at. Since TEL bases are known even to open-source analysts, we give the US credit for "gatekeeping" all TEL bases, monitoring entry and exit traffic.[345]

**Roads**
US intelligence may also have an understanding, built up over years, of China's typical TEL patrol routes. It almost certainly understands which parts of China constitute drivable terrain for TELs, and which roads and/or fields are too steep, soft, or insecure to admit TEL traffic – in other words, using pure road density

[341] Glaser and Fetter, "Correspondence," 205-6.

[342] Wu, "Living with Uncertainty," 118.

[343] Long and Green, "Stalking the Secure Second Strike".

[344] Glaser and Fetter, 2018.

[345] Alan J. Vick, Richard M. Moore, Bruce R. Pirnie, and John Stillion, "Aerospace Operations Against Elusive Ground Targets," *RAND*, 2001, https://www.rand.org/pubs/monograph_reports/MR1398.html.

overestimates drivable TEL area and represents a conservative estimate, a ceiling for where TELs could possibly drive). This effectively decreases the search area for broad collection modalities substantially.

**TELs**
Finally, the US likely has an intimate understanding of each TEL type's height, weight, and so on. In fact, precise specifications for TELs are available even in the open-source. For example, we know the length, diameter, and mass of the DF-31 missile, canister, and associated TEL vehicle.[346]

This knowledge also gives the US knowledge of TEL-specific signatures. In particular, since China has itself paraded these missiles through its streets, and each TEL type was in development for many years and has existed for years after that, it seems safe to assume that the US has an arbitrary quantity of images of each TEL type (one can find many images on the open-source Internet), making ML training possible.

Furter, the US knows TELs have various signatures that make them quite unlike civilian trucks. In addition to having a particular mass, shape, and brute appearance, mobile solid-propellant missiles require "approximately six vehicles, including the TEL."[347] These include service trucks purposed for fire control, power, power distribution, aiming, and inspection.[348]

## (4) US Intelligence: Broad-Area Detection

We give the US credit for three kinds of broad-area detection assets: EO satellites, SAR/GMTI satellites, and standoff aircraft.

**EO satellites**
Mature miniaturization of electro-optical satellites likely means continuous high-resolution imaging from space is available to the Pentagon for every part of the Earth.[349] EO satellites, of course, can be temporarily defeated by bad weather and cloud cover; their view can also be blocked by high valleys or buildings. Based on Trump's tweeting of a KH-11 image, US EO satellites have at least 10-cm resolution.[350]

**SAR/GMTI satellites**
Recent technological advances mean satellites using synthetic aperture radar can detect both stationary and moving targets at sufficient resolution to find truck-sized objects.[351] SAR satellites in sun-synchronous low earth orbit (LEO) can iteratively view the entire earth, but consequently provide only intermittent coverage of any given area. The United States has a finite number of SAR satellites, but ongoing miniaturization likely will mean progressively reduced time between usable passes.[352] SAR satellites provide day-night all-weather coverage, though they can be blocked by disadvantageous topography.

---

[346] Li, "Tracking Chinese Strategic Mobile Missiles."

[347] "New Mobile Solid-Propellant MRBM Under Development, China (S)," *Central Intelligence Agency*, November 1983, https://www.cia.gov/library/readingroom/docs/CIA-RDP84T00171R000300410001-4.pdf, 8.

[348] Wu, "Certainty of Uncertainty," 586.

[349] Glaser and Fetter, "Should the United States Reject MAD?", 69; Glaser and Fetter, "Correspondence," 205.

[350] Daniel Oberhaus, "Trump Tweeted a Sensitive Photo. Internet Sleuths Decoded It," *WIRED*, September 3, 2019, https://www.wired.com/story/trump-tweeted-a-sensitive-photo-internet-sleuths-decoded-it/.

[351] Lieber and Press, "The New Era of Counterforce."

[352] For example, the Finnish SAR satellite company ICEYE has the ambitious motto "every meter, every hour."

SAR miniaturization means the possibility of "near-continuous all-weather day-night coverage" at progressively cheaper cost.[353] Besides relying on military satellites, the rapid proliferation of commercial constellations due to falling launch costs and ongoing miniaturization also provides a ready-made source of data. These commercial constellations would complicate Chinese counterspace efforts, which would need to destroy hundreds of distributed assets belonging to third parties to disable their use; this would represent both a targeting problem and a political one.[354] Benchmarking US capabilities against commercial SAR provider ICEYE, the US arguably should have the ability to view each meter of China once per hour.

**Standoff aircraft**

Finally, aircraft carrying SAR operating in the GMTI mode can also both detect and track Chinese TELs. One way to do this would be to operate off the Chinese coast ("standoff"). US aircraft regularly fly reconnaissance mentions close to the Chinese coast, including sometimes disguised as aircraft from other countries (e.g., Malaysia, the Philippines).[355] There were 60 such flights in September, for example.[356] These include the US Navy's P-8A Poseidon planes, equipped with AN/APY-10 radars which can conduct surface searches with SAR, ISAR, and also take simple video from up to 370 km away. It also includes the EP-3E ARIES II, which also carries AN/APY-10 series radar.[357] The EP-3E ARIES II also carries ELINT and COMINT systems with a maximum range of 926 km. All together, US standoff aircraft can be understood as likely seeing about 400 km inwards from the coast with both video and SAR/ISAR/GMTI day-night all-weather radar, and with ELINT/COMINT systems about 900 km.[358]

## (5) US Intelligence: Targeted Collection

**Base surveillance**

TEL bases themselves do not migrate. Thus, the US can pre-position intelligence assets to gate-keep these bases, more effectively observing when TELs enter and leave the area.

First, consider unattended ground sensors. The US could position UGSs near bases and major roads TELs are likely to turn onto, allowing them to distinguish between TELs and lighter traffic.[359] UGSs with decade-long lifespans appear possible. After emplacement by UAV or human agents, they could remain silent until

[353] Charles L. Glaser and Steve Fetter, "Correspondence: The Limits of Damage Limitation," *International Security* 42.1 (2017), 205.

[354] Matthew A. Hallex and Travis S. Cottom, "Proliferated Commercial Satellite Constellations: Implications for National Security," *Joint Forces Quarterly* 97 (2020), 24. Available online: https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-97/jfq-97_20-29_Hallex-Cottom.pdf?ver=2020-03-31-130614-940.

[355] Liu Zhen, "US puts record number of eyes in the skies off China coast in July," *South China Morning Post*, July 25, 2020.

[356] Ralph Jennings, "China Reports Spike in US Surveillance Flights," *VOA News*, October 23, 2020, https://www.voanews.com/east-asia-pacific/voa-news-china/china-reports-spike-us-surveillance-flights.

[357] John Pike, "AN/APY-10 Multi-Mission Maritime and Overland Surveillance Radar," *Global Security*, December 12, 2020, https://www.globalsecurity.org/military/systems/aircraft/systems/an-apy-10.htm; John Pike, "EP-3E ARIES," *Global Security*, July 28, 2011, https://www.globalsecurity.org/intell/systems/ep-3.htm; Steven Aftergood, "EP-3E ARIES," October 20, 2016, https://fas.org/irp/program/collect/ep-3_aries.htm.

[358] Lieber and Press, "The New Era of Counterforce"; Green and Long, "Correspondence."

[359] Green and Long, "Stalking the Secure Second Strike."

activated, recharging via sunlight.[360] UGSs could also help cue any penetrating UAVs.[361] Indeed, DARPA has studied using deep learning to fuse and class signatures derived from physical sensors.[362]

Second, consider geosynchronous SIGINT satellites. The US historically tracked Soviet TELs in part by intercepting communications using SIGINT satellites.[363] The US continues to develop these capabilities today, with continuous launches of more SIGINT satellites. Cell phone, radio, and camera intercepts could enable "listen[ing] into the conversations of technicians working in a missile convoy."[364] It is very hard for TELs not to talk, compared to fixed missile sites.[365] While the details are of course classified, reporting seems to indicate that the US Orion constellation can listen to most cell communications in China at most times. Once cued to stare at the right places, the US can listen to 100% of TELs that are not under EMCON.[366] While this traffic is likely encrypted, we estimate based on broad technical assumptions that the US has a 5% chance of bypassing encryption for any given intercept.

**Cueing**

Once detected by broad-area assets, others which have narrower range can be "cued" to migrate toward the identified TELs, providing additional modalities. First, hyperspectral collectors can powerfully distinguish between decoys and real TELs.[367] There is a limited number of these, but their numbers are increasing.[368]

Second, although higher-risk, the US could send stealthy high-endurance unmanned drones into Chinese airspace, likely launched through non-coastal parts of the Chinese border (e.g., through Pakistan). This modality is new with high-endurance UAVs, which unlike human pilots are happy to crawl through airspace for very long periods. Stealthy UAVs are thus useful for tracking cued targets.[369]

---

[360] Noah Shachtman, "This Rock Could Spy on You for Decades," *WIRED*, May 29, 2012, https://www.wired.com/2012/05/spy-rock/. Mark Hewish, "Reformatting fighter tactics," *Jane's International Defense Review*, 2007. Available online: https://web.archive.org/web/20070815215345/http://www.textrondefense.com/pdfs/news/jidr06_01.pdf.

[361] Stephen C. Stubberud and Kathleen A. Kramer, "UAVs working in conjunction with unattended ground sensors," *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*.

[362] Benjamin Epstein and Roy H. Olsson III, "Physical Signal Classification Via Deep Neural Networks," *arXiv*, November 15, 2018, https://arxiv.org/abs/1811.06349.

[363] A. Andronov, "American signals intelligence satellites in geosynchronous orbit," *Foreign Military Review* 12, trans. Allen Thomson (1993), 42; Long and Green, "Stalking the Secure Second Strike," 51–56, 60–64.

[364] Paul Bracken, "The Intersection of Cyber and Nuclear War," *Strategy Bridge*, January 17, 2017, https://thestrategybridge.org/the-bridge/2017/1/17/the-intersection-of-cyber-and-nuclear-war.

[365] Glaser and Fetter, "Correspondence"; Bracken, ibid.

[366] Zulfikar Abbany, "Modern spy satellites in an age of space wars," *Deutsche Welle*, August 25, 2020, https://www.dw.com/en/modern-spy-satellites-in-an-age-of-space-wars/a-54691887; Andronov, "American signals intelligence satellites." Marco Langbroek, "A NEMESIS in the sky," *The Space Review*, October 31, 2016, https://www.thespacereview.com/article/3095/1. Reporting based in part on Intercept leaks of US documents.

[367] Vick et al., "New Concepts for Defeating Mobile Missiles," 71, 76.

[368] Adam Keith, "Is hyperspectral the next Earth observation frontier?," *SpaceNews*, March 30, 2019, https://spacenews.com/op-ed-is-hyperspectral-the-next-earth-observation-frontier/.

[369] Lieber and Press, "The New Era of Counterforce," 43-6; Green and Long, "Correspondence."

We give the US credit for 20 RQ-180s, identical to the number of actively fielded RQ-170s. The RQ-180 reportedly has parameters similar to the Global Hawk, giving it 24 hours of flight time, 629 km/h speed, and a range of 22,000 km; it is reportedly even stealthier than the F-35. As Heginbotham et al. (2015) found that latest-generation aircraft with very low detectability could still access 93% of a notional target set in China in 2017, we give the RQ-180 a significant probability of penetration.[370] We assume the RQ-18 has a maintenance to flight time ratio of 1:1, and model it as a source of EO and SAR/GMTI data.

### (6) US Intelligence: AI-Assisted Processing

To quantitatively estimate AI's effects, we stipulate that satellite data arrives as 250m x 250m tiles, since our search is for small trucks and we have access to at least 1m-resolution imagery. Consequently, for every square kilometer, we have 16 images to consider.

What about false positives? Naively, China has an area of around 10 million sq km, so we would need 160 million images to cover the entire country. Obviously, however, not all areas are near TEL bases, some areas definitely do not contain TELs (e.g., inaccessible mountaintops), and our simulation assumes TELs drive on roads. Adjusting for China's road network to bound uncertainty, we find the need to process the equivalent of between 37,500 and 1,250,000 sq km. How long would this take?

Happily, a study by Marcum et al. (2017) analyzing the kindred problem of broad-area search for SAM sites took 13 hours to process 88,000 sq km of data, using one server and 4 GPUs.[371] This computation is embarrassingly parallel, however, so speed can be trivially scaled by adding additional computing power. If we assume the US government uses computation equivalent to machine learning applications for one Google datacenter, this would roughly generate 1,000 times the computing power of Marcum et al.'s setup, and therefore achieve about 100 times lower latency while processing an order of magnitude more data. Consequently, accounting for government willingness to spend what it takes, as well as optimizations specific to satellite imagery, we estimate a processing time of 5 minutes.

Marcum et al. (2017) also usefully provides parameters for expert human analysts: in their experiment, the processing task required 43 minutes, with human analysts able to review an image every 1.3 seconds using an optimized user interface. Assuming the United States could devote 100 analysts to a counterforce effort, this means our notional human analysis team processes 7800 instances per minute. Consequently, AI represents an enormous increase over human-only efficiency. With human-machine teaming, we set algorithmic false positives and false negatives locally based on truck density, then pass on the most controversial images to the human team.

---

[370] Heginbotham et al., "The U.S.-China Military Scorecard," 112. Guy Norris, "USAF Unit Moves Reveal Clues To RQ-180 Ops Debut," *Aviation Week & Space Technology*, October 23, 2019, https://aviationweek.com/defense-space/usaf-unit-moves-reveal-clues-rq-180-ops-debut; Amy Butler and Bill Sweetman, "Secret New UAS Shows Stealth, Efficiency Advances," *Aviation Week*, December 6, 2013; Jen DiMascio, "Unmasking the RQ-180," *Aviation Week & Space Technology*, December 6, 2013.

On stealth techniques used by the immediate predecessor, the RQ-170, see David Axe, "7 Secret Ways America's Stealth Armada Stays Off the Radar," *WIRED*, 2012, https://www.wired.com/2012/12/steath-secrets/?pid=1688. Some RQ-170 details were revealed after one crashed while monitoring Iranian nuclear activity. See Scott Shane and David E. Sanger, "Drone Crash in Iran Reveals Secret U.S. Surveillance Effort," *The New York Times*, December 7, 2011, https://www.nytimes.com/2011/12/08/world/middleeast/drone-crash-in-iran-reveals-secret-us-surveillance-bid.html.

[371] Richard A. Marcum, Curt H. Davis, Grant J. Scott, and Tyler W. Nivin, "Rapid broad area search and detection of Chinese surface-to-air missile sites using deep convolutional neural networks," *Journal of Applied Remote Sensing* 11.4 (2017).

## (7) US Counterforce Against Chinese TELs

How would US nuclear counterforce against Chinese TELs work? We first enumerate the US nuclear arsenal, then discuss specifics involved in destroying TELs.

### US Nuclear Forces

The United States possesses approximately 3800 nuclear warheads, of which 1750 are actively deployed: 400 on land-based intercontinental ballistic missiles (ICBMs), 900 on submarine-launched ballistic missiles (SLBMs), 300 at bomber bases on US soil, and 150 tactical bombs at European bases.[372] We consider only SLBMs and ICBMs, given that long aircraft travel times would not be suitable for striking mobile targets with fleeting locations. Since China does not possess effective ballistic missile defenses, the destroyable area calculations can be performed cleanly. Further, based on estimating how many fixed structures the US would seek to destroy specifically with nuclear weapons, we reserve 240 ICBMs (180 W78s, and 60 W87s) for that purpose. As a sanity check, Heginbotham et al. (2015) estimate 157 US warheads for fixed targets in their 2017 US first-strike scenario.[373]

### Available TEL-Killing Weapons

| Weapon | Available Count |
|---|---|
| W76-1 | 675 |
| W88 | 225 |
| W78 | 120 |
| W87 | 40 |

**SLBMs**

The US possesses 1,486 UGM-133A Trident II D5/LEs ("Trident D5s"), with 900 to 950 warheads normally deployed on submarines at any given time. Since Trident D5s possess an extended range of 12,000 km, US submarines can strike all of China from even just off the west coast. Further, more than 60 percent of deterrent patrols take place in the Pacific, likely reducing SLBM travel time by half or more.[374] Trident D5s travel at 29,020 km/h; consequently, SLBMs fired from maritime East Asia will have flight times of 12.4 minutes or less, while those outside will have maximum travel times of 24.8 minutes. We assume the deployed mixture consists of 75% Mk4As and 25% Mk5s, which matches the proportion of available warheads in the US arsenal.

**ICBMs**

The US maintains 400 silo-based LGM-30G Minuteman III ICBMs deployed across Colorado, Nebraska, Wyoming, North Dakota, and Montana. Consequently, given a speed of 24,000 km/h, maximum flight times to China are half an hour.[375] We similarly assume the deployed mixture consists of 75% Mk12As and and 25% Mk21/SERVs, which matches the proportion of available warheads in the US arsenal.

---

[372] Hans M. Kristensen and Matt Korda, "United States nuclear forces, 2020," *Bulletin of the Atomic Scientists* 76.1 (2020): 46-60.

[373] Heginbotham et al., "The U.S.-China Military Scorecard," 314.

[374] Kristensen and Korda, "United States nuclear forces, 2020," 47, 52-4.

[375] "LGM-30 Minuteman III," *Federation of American Scientists*, October 20, 2016, https://fas.org/nuke/guide/usa/icbm/lgm-30_3.htm.

## TEL-Killing

To set parameters for TEL-killing, we must estimate both how large an area must be destroyed at what psi, as well as how nuclear weapon yields convert to area.

### Target Sizing

The target area which must be destroyed to neutralize a TEL depends on three factors: TEL speed, road density, and weapon flight time. On speed, since TELs derive survivability partially through resembling surrounding non-TEL traffic, TELs likely drive at speeds normal for Chinese roads. Following Chinese road regulations, this reportedly limits TEL speeds to between 20 km/h and 40 km/h, although they could hypothetically accelerate to 69 km/h, the maximum speed of the DF-31 TEL.[376] On road density, Chinese road density in 2019 was 52.21 km per 100 sq km, though area information unfortunately appears unavailable.[377] If TELs could drive in any direction, the area the TEL could be after *n* minutes would obviously be a circle, $\pi r^2$. However, Wu (2020) claims that "Chinese TELs lack off-road mobility."[378] We assume fully half of China is TEL-drivable, which likely overstates the case significantly; thus, half that area must be destroyed.

Finally, weapon flight time ("delay") will obviously vary by platform. As above, we give the United States credit for locating 60% of its SSBNs in the Pacific, with maximum flight times of 12.4 minutes; the remaining 40% have flight times of up to 24.8 minutes. CONUS-based ICBMs experience flight times of up to half an hour. Assuming TELs move during weapon flight, neutralizing any given TEL requires striking the whole area that TEL could be, after weapon arrival. Thus, TELs represent a target area of:

$$\textbf{road density} * \pi * (\textbf{speed} * \textbf{delay})^2$$

where road density is set to 0.5, speed varies according to alert level as above, and delay varies by weapon.

### Strike Area

How much area does each weapon strike cover? This depends on two primary factors: TEL hardness and atack doctrine. On hardness, TELs can be neutralized by 2-4 psi; we conservatively round up to 5 psi.[379] On attack doctrine, we follow Lieber and Press (2017) in their assessment that with modern weapons technology, two nuclear weapons should destroy a targeted area with a probability of kill (Pk) exceeding 99.9%.[380] Consequently, we assume the US follows a 2:1 strike doctrine for TELs, seeking to blanket the area possibility containing the TEL with 5 psi, twice.

We model the relationship between each weapon and the area its detonation covers using two sources: first, we use the surface burst formula commonly cited by Li (2007), Glaser and Fetter (2016), Lieber and Press

[376] Bin, "Tracking Chinese Strategic Mobile Missiles," 9-10; Genys, "Hanyang HY4330."

[377] Samantha Wong, "Road density in China 2008-2019," *Statista*, November 24, 2020, https://www.statista.com/statistics/258345/road-density-in-china/.

[378] Wu Riqiang, Appendix to "Living with Uncertainty: Modeling China's Nuclear Survivability," Harvard Dataverse, 2020, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5EKNJM, 4.

[379] Daryl Press, "Simple Mathematics of Nuclear Force Analysis" (presentation, Summer Nuclear Bootcamp, Albuquerque, New Mexico, August 2019).

[380] Lieber and Press, "The New Era of Counterforce," 26.

(2017), and Wu (2020). This formula ultimately derives from the standard reference text by Glasstone and Dolan, *Effects of Nuclear Weapons* (1977):[381]

$$LR = \frac{2.45Y^{1/3}}{H^{1/3}} \left\{ \sqrt{1 + \frac{2.79}{H}} + \frac{1.67}{H^{1/2}} \right\}^{2/3}$$

Here, LR is lethal radius in nautical miles, Y is yield in megatons, and H is target hardness in psi. However, this arguably underestimates the efficacy of the US nuclear arsenal against TELs, as soft targets do not require surface burst detonations – an airburst is likely to produce sufficient force across a wider area.

Thus, second, we source airburst data from Alex Wellerstein's NukeMap utility, which also implements its formula and graphs from *Effects of Nuclear Weapons*.[382] Our calculations agree approximately with the US Defense Intelligence Agency's 1969 *Physical Vulnerability Handbook – Nuclear Weapon*s, which holds that a 100-kt weapon can effectively destroy road-mobile missiles across a radius of 2875 meters, or approximately 26 square kilometers.[383] This produces an increase of the efficacy of the US nuclear arsenal against 5 psi TELs by very roughly a factor of 3.

**TEL-Killing Area by Weapon**

| Weapon | Surface Burst (sq km) | Airburst (sq km) |
|---|---|---|
| W76-1 | 11.16 | 31.2 |
| W88 | 32.86 | 91.9 |
| W78 | 26.80 | 74.9 |
| W87 | 24.90 | 69.6 |

Consequently, we can solve for a US first-strike package by combining TEL target sizing and weapon strike area, keeping in mind the different delays for different weapons. In tables:

**TEL-Killing Data (Pacific SLBMs)**

| TEL Speed (km/h) | Average Road Density, 2019 (km / sq km) | Pacific SLBM Delay (h) | Target Area (sq km) | W76-1s Needed (Surface Burst, 2:1) | W76-1s Needed (Airburst, 2:1) | W-88s Needed (Surface Burst, 2:1) | W-88s Needed (Airburst, 2:1) |
|---|---|---|---|---|---|---|---|
| 20 | 0.5221 | 0.207 | 28.11 | 5.04 | 1.80 | 1.71 | 0.61 |
| 30 | 0.5221 | 0.207 | 63.25 | 11.34 | 4.05 | 3.85 | 1.38 |

[381] Lynn Etheridge Davis and Warner R. Schilling, "All You Ever Wanted to Know about MIRV and ICBM Calculations but Were Not Cleared to Ask," *The Journal of Conflict Resolution* 17.2 (1973), 213.

[382] For an explanation of airburst calculations, see Alex Wellerstein, "The trouble with airbursts," Restricted Data, December 6, 2013, http://blog.nuclearsecrecy.com/2013/12/06/trouble-airbursts/.

[383] Matthew G. McKinzie, Thomas B. Cochran, Robert S. Norris, and William M. Arkin, "The U.S. Nuclear War Plan: A Time for Change," *Natural Resources Defense Council*, 2001, https://www.nrdc.org/sites/default/files/us-nuclear-war-plan-report.pdf, 54.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 40 | 0.5221 | 0.207 | 112.45 | 20.15 | 7.21 | 6.84 | 2.45 |
| 50 | 0.5221 | 0.207 | 175.71 | 31.49 | 11.26 | 10.69 | 3.82 |
| 60 | 0.5221 | 0.207 | 253.02 | 45.34 | 16.22 | 15.40 | 5.51 |
| 70 | 0.5221 | 0.207 | 344.38 | 61.72 | 22.08 | 20.96 | 7.49 |
| 80 | 0.5221 | 0.207 | 449.80 | 80.61 | 28.83 | 27.38 | 9.79 |
| 90 | 0.5221 | 0.207 | 569.28 | 102.02 | 36.49 | 34.65 | 12.39 |
| 100 | 0.5221 | 0.207 | 702.82 | 125.95 | 45.05 | 42.78 | 15.30 |

**TEL-Killing Data (Distant SLBMs)**

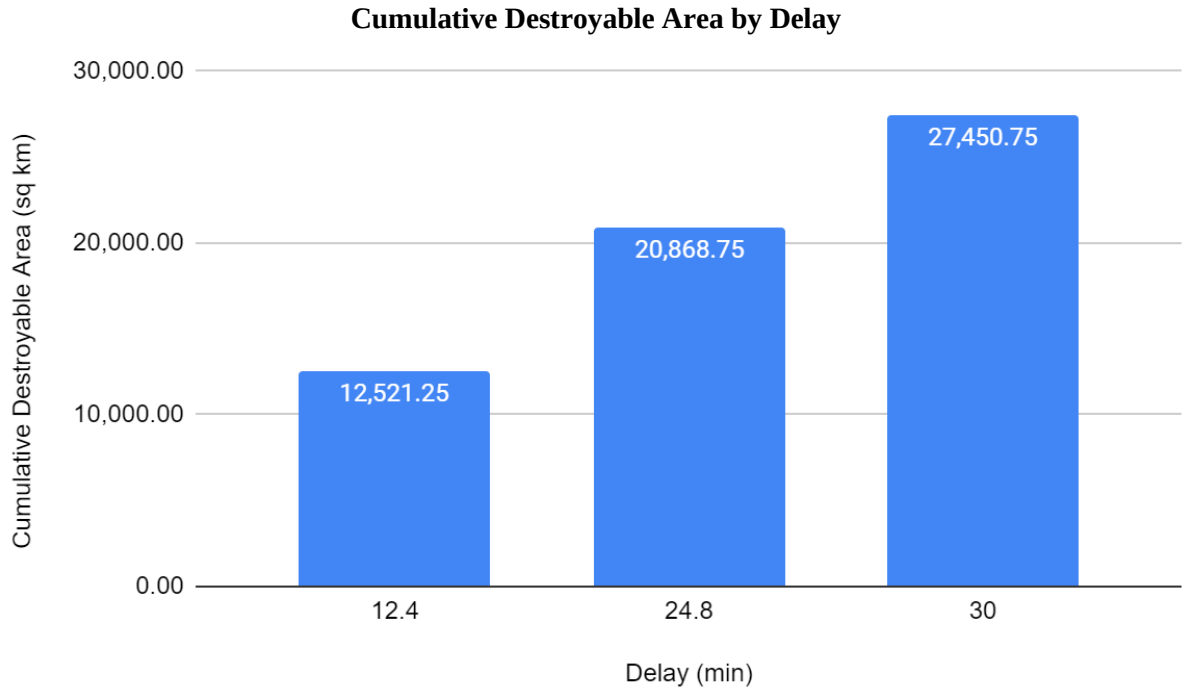
| TEL Speed (km/h) | Average Road Density, 2019 (km / sq km) | Distant SLBM Delay (h) | Target Area (sq km) | W76-1s Needed (Surface Burst, 2:1) | W76-1s Needed (Airburst, 2:1) | W-88s Needed (Surface Burst, 2:1) | W-88s Needed (Airburst, 2:1) |
|---|---|---|---|---|---|---|---|
| 20 | 0.5221 | 0.413 | 111.91 | 20.06 | 7.17 | 6.81 | 2.44 |
| 30 | 0.5221 | 0.413 | 251.79 | 45.12 | 16.14 | 15.33 | 5.48 |
| 40 | 0.5221 | 0.413 | 447.63 | 80.22 | 28.69 | 27.24 | 9.74 |
| 50 | 0.5221 | 0.413 | 699.43 | 125.35 | 44.84 | 42.57 | 15.22 |
| 60 | 0.5221 | 0.413 | 1,007.18 | 180.50 | 64.56 | 61.30 | 21.92 |
| 70 | 0.5221 | 0.413 | 1,370.88 | 245.68 | 87.88 | 83.44 | 29.83 |
| 80 | 0.5221 | 0.413 | 1,790.54 | 320.89 | 114.78 | 108.98 | 38.97 |
| 90 | 0.5221 | 0.413 | 2,266.15 | 406.12 | 145.27 | 137.93 | 49.32 |
| 100 | 0.5221 | 0.413 | 2,797.72 | 501.38 | 179.34 | 170.28 | 60.89 |

**TEL-Killing Data (ICBMs)**

| TEL Speed (km/h) | Average Road Density, 2019 (km / sq km) | CONUS ICBM Delay (h) | Target Area (sq km) | W78s Needed (Surface Burst, 2:1) | W78s Needed (Airburst, 2:1) | W-87s Needed (Surface Burst, 2:1) | W-87s Needed (Airburst, 2:1) |
|---|---|---|---|---|---|---|---|
| 20 | 0.5221 | 0.5 | 164.02 | 12.24 | 4.38 | 13.17 | 4.71 |
| 30 | 0.5221 | 0.5 | 369.05 | 27.54 | 9.85 | 29.64 | 10.60 |
| 40 | 0.5221 | 0.5 | 656.09 | 48.96 | 17.52 | 52.70 | 18.85 |
| 50 | 0.5221 | 0.5 | 1,025.14 | 76.50 | 27.37 | 82.34 | 29.46 |
| 60 | 0.5221 | 0.5 | 1,476.20 | 110.16 | 39.42 | 118.57 | 42.42 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 70 | 0.5221 | 0.5 | 2,009.28 | 149.95 | 53.65 | 161.39 | 57.74 |
| 80 | 0.5221 | 0.5 | 2,624.36 | 195.85 | 70.08 | 210.79 | 75.41 |
| 90 | 0.5221 | 0.5 | 3,321.46 | 247.87 | 88.69 | 266.78 | 95.44 |
| 100 | 0.5221 | 0.5 | 4,100.56 | 306.01 | 109.49 | 329.36 | 117.83 |

For visual legibility, we can also graph the cumulative destroyable area (within which TELs hardened to 5 psi would be destroyed) against delay from weapon launch. To accomplish this, we simplify multiply through weapon availability by the 2:1 airburst tables given above, sorting by delay. Following these assumptions, we estimate that the United States can destroy 27,450.75 sq km of TELs after 30 minutes.




**What about tunnels?**

Our simulation does not explicitly model tunnels, but at minimum cave-ins of all exits could be easily accomplished with even non-nuclear strikes; depending on the facility in question, surface-burst attacks with nuclear weapons should also penetrate immense quantities of earth. According to a Chinese study, US nuclear weaponry "would thoroughly destroy the tunnel exits/entrances of China's underground missile sites."[384] Consequently, TELs in tunnels arguably resemble fixed targets.

**What about launchpads?**

Strangely, Wu (2020) assumes that Chinese mobile missiles are each assigned to three corresponding pre-surveyed launch sites, and that the missiles will be unable to launch if those sites are destroyed.[385] As Wu

[384] Wu, "Living with Uncertainty," 101. Eric M. Sepp, "Deeply Buried Facilities: Implications for Military Operations," *Occasional Paper* 14 (2000), 25, 31.

[385] Ibid., 90-1.

himself notes, however, there is not robust evidence this is true. Indeed, there is no obvious technical reason why it would be thus. Soviet launchers did not require pre-prepared sites, so it is at least possible China will develop the capability to launch from arbitrary areas meeting certain minimum conditions.[386] We assume TELs can launch from arbitrary locations, and require the United States to destroy the TEL itself.

**What about deserts?**

Glaser and Fetter (2018) suggest China could deploy TELs throughout the Gobi Desert, which contains an area larger than Texas mostly comprised of bare rock. This would enable omnidirectional maneuvering at high speeds, potentially overwhelming the US nuclear arsenal by presenting an extremely large effective target size.[387]

However, this strategy would not be costless, as the nature of the Gobi Desert would complicate TEL concealment. Deserts are mostly cloudless and experience little rainfall, meaning TELs could not leverage periods of bad weather to relocate without being imaged by US EO satellites. An environment of mostly bare rock would also prevent TELs from using topography to avoid US SAR satellite passes. With little of anything else around, TELs would be easily distinguishable against the featureless ground; rather than rely on wide-area nuclear barrages, the United States could bring conventional assets to bear on these wayward nomads, especially given the Gobi Desert's presumably low density of air defense capabilities.

Further, more importantly, Glaser and Fetter credit these Gobi TELs with various technological improvements, such as traveling at high-speeds off-road and being hardened to resist 5-10 psi (5 psi is generally understood to suffice for destroying cities). While these are plausible, the US would naturally develop appropriate countermeasures; one can also arbitrarily make the US arsenal sufficient again to destroy such TELs, for example, by simply increasing the delivery speed and/or yield of US nuclear weapons to generate a sufficient area of destruction. Unless some new technical development changes matters, in other words, it is always possible to make any arsenal survivable or vulnerable simply by stipulating that the other side fails to respond to increased arming. In contrast, AI's effects are notable because they leverage the effects of already-existing intelligence assets, acting as a force multiplier.

## Other Assumptions

Finally, it is worth noting that in modeling the US counterforce effort, we have made a number of simplifying assumptions which generally favor China. First, besides destroying the launchers themselves, the United States could also preclude a Chinese second strike by disabling its command and control. We do not explore this possibility here, though this would likely be easier than destroying each mobile Chinese TEL, as most command-and-control assets are fixed.

Second, the United States could also leverage conventional assets for counterforce. Large-area barrage attacks against TELs demand the large blast radii of nuclear weapons, but prepared launch sites, special railroads designated for transporting warheads, visible shelters, and other targets could all be destroyed with various sources of long-range precision strike. Finally, we assume the US cannot distinguish between mated and de-mated TELs in the field, and therefore must destroy both. In practice, various intelligence sources would yield some hope of discrimination.

[386] Matthew G. McKinzie, Thomas B. Cochran, Robert S. Norris, and William M. Arkin, "The U.S. Nuclear War Plan: A Time for Change," *Natural Resources Defense Council*, 2001, https://www.nrdc.org/sites/default/files/us-nuclear-war-plan-report.pdf, 52. Soviet road-mobile missiles could launch from unprepared sites if "terrain relief" allowed.

[387] Glaser and Fetter, "Correspondence," 205-6.

## (8) Missile Defense Against Surviving TELs

We implement an extremely simple Monte Carlo layer to account for US missile defenses: giving the US credit for its 44 interceptors, we assume each has Pk of 0.5, and that all are fired simultaneously, distributed equally, against any approaching Chinese missiles. For simplicity, we make the China-favoring assumption of giving the PLARF credit for perfect weapon reliability of any surviving TELs. Since the probability that one missile survives one interceptor is (1 – Pk), the probability that one missile survives its equal share interceptors is:

$$(1 - Pk)^{\frac{N}{M}}$$

where N is the number of interceptors, and M is the number of incoming missiles. Thus, the probability that the missile is destroyed is

$$1 - (1 - Pk)^{\frac{N}{M}}$$

and the probability that M missiles are destroyed is, correspondingly,

$$(1 - (1 - Pk)^{(\frac{N}{M})})^{M}$$

Summing up, the probability that at least one missile is not destroyed is then

$$1 - (1 - (1 - Pk)^{(\frac{N}{M})})^{M}$$

We can also produce this, practically, in table form, assuming the US would be deterred by even one hit:

**US Missile Defense Efficacy**

| Surviving Missiles | Probability of ≥ 1 Hit | Surviving Missiles | Probability of ≥ 1 Hit |
|---|---|---|---|
| 1 | 0.00% | 13 | 72.98% |
| 2 | 0.00% | 14 | 81.40% |
| 3 | 0.01% | 15 | 87.81% |
| 4 | 0.20% | 16 | 92.38% |
| 5 | 1.12% | 17 | 95.46% |
| 6 | 3.66% | 18 | 97.41% |
| 7 | 8.63% | 19 | 98.59% |
| 8 | 16.37% | 20 | 99.26% |
| 9 | 26.58% | 21 | 99.63% |
| 10 | 38.45% | 22 | 99.82% |
| 11 | 50.83% | 23 | 99.92% |
| 12 | 62.63% | 24 | 99.96% |
| 13 | 72.98% | 25 | 99.98% |
| 14 | 81.40% | 26 | 99.99% |

## Results

Overall, we find evidence for the pessimistic view in the ongoing scholarly discussion of the possibility of US-China nuclear war due to strategic instability. Under conditions of low alert, we find US detection of Chinese road-mobile missiles is occasionally possible even without employing AI: windows of opportunity periodically emerge. With AI, we find the United States can *consistently* track Chinese TELs in the low alert case, and that windows of opportunity emerge even under higher levels of alert.

We run various world-states through low, medium, and high alert simulations for twenty-four hours each, beginning with the base case, using best-guess point estimates plus AI, and the base case minus AI; we then discuss some simulation excursions, and then potential countermeasures by states concerned about AI.

### (1) Base Case: Nuclear Counterforce, With and Without AI, All Alert Levels

We begin by visualizing the comparison of the two base cases, with and without AI, on the next page. Throughout, we focus primarily on four metrics:

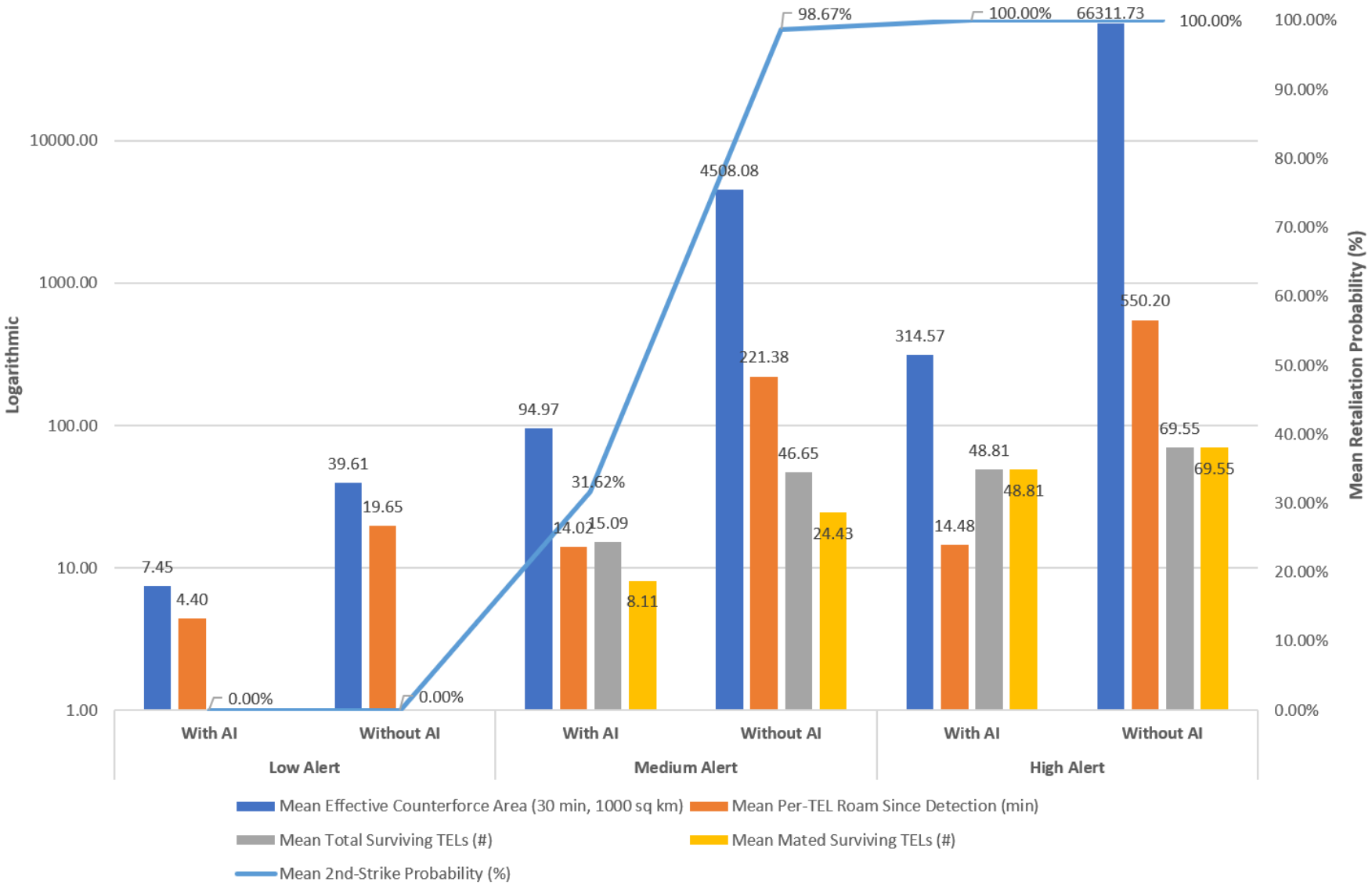
- **undetected roaming,** or the average number of minutes spent roaming since the average TEL was last tracked by the United States. This provides a measure of TEL elusiveness over the lifetime of the simulation; larger delays imply increasingly imprecise US understandings of TEL locations.

- **the effective counterforce area,** or the total area in square kilometers that the United States must destroy, at each time-step in the simulation, to successfully execute a first-strike against China. Since this area depends on weapon latency, we graph three lines corresponding to average flight times for US SLBMs within and without East Asia, as well as ICBMs originating from CONUS. As explained above, the counterforce area depends primarily on TEL speed and US knowledge of TEL locations.

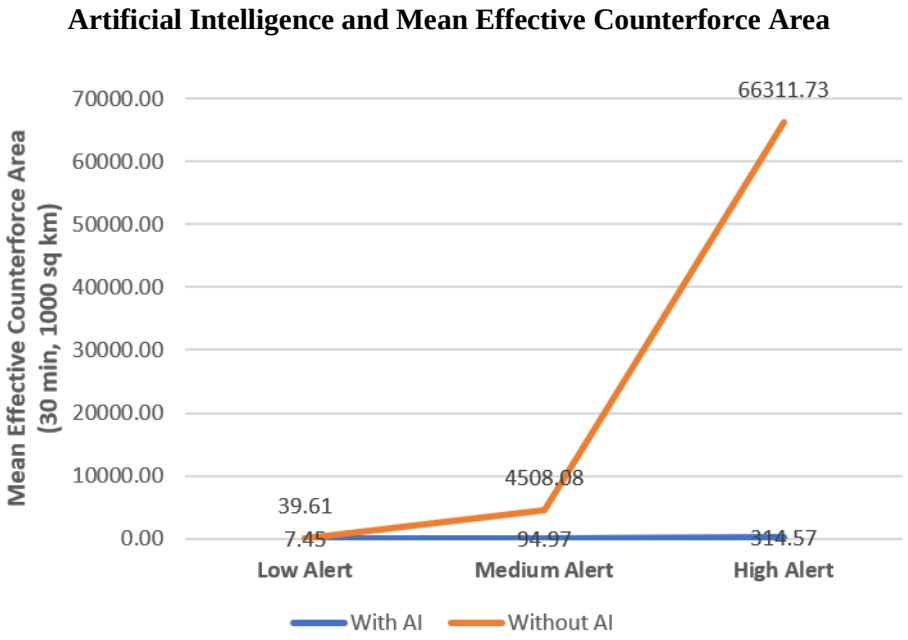
- **surviving TELs,** meaning how many TELs would remain if the United States initiated counterforce efforts at each time-step. Note that since not all TELs are mated under low and medium alert and the US possesses missile defenses, this is not identical to whether China has successfully retained a second-strike capability.

- **second-strike probability,** or the chance that China would successfully strike US soil with at least one penetrating missile after absorbing the US counterforce attack if all surviving mated TELs fired, taking into account US missile defenses.

**Base Case: Nuclear Counterforce, With and Without AI, All Alert Levels**



First, most notably, AI's large effect size above is clearly evident, so much so that a logarithmic scale became necessary. To illustrate this more clearly, we isolate AI from above and graph it non-logarithmically:

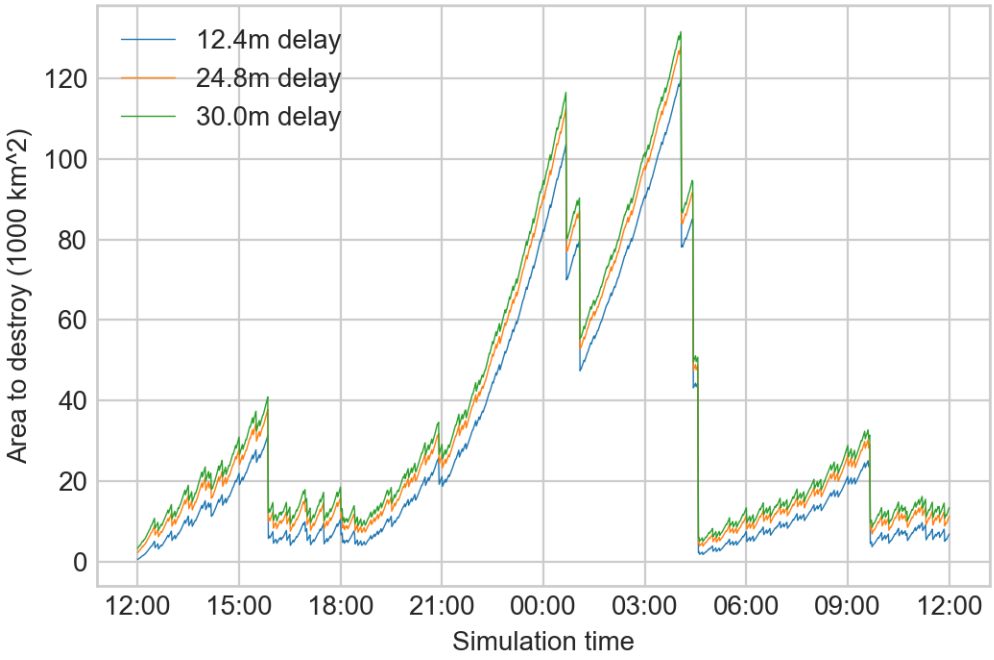



In the mechanics of our simulation, AI contributes immensely to US counterforce efforts, making the area which must be searched by other modalities or subject to barrage attack less by two entire orders of magnitude, in both the high and medium alert cases, and one order of magnitude in the low alert case. Equivalently, the effect of AI-powered intelligence-processing, in terms of US ability in a counterforce effort, has an effect size in our model equivalent to increasing the brute size of the US arsenal by between 1,000% and 100,000%.

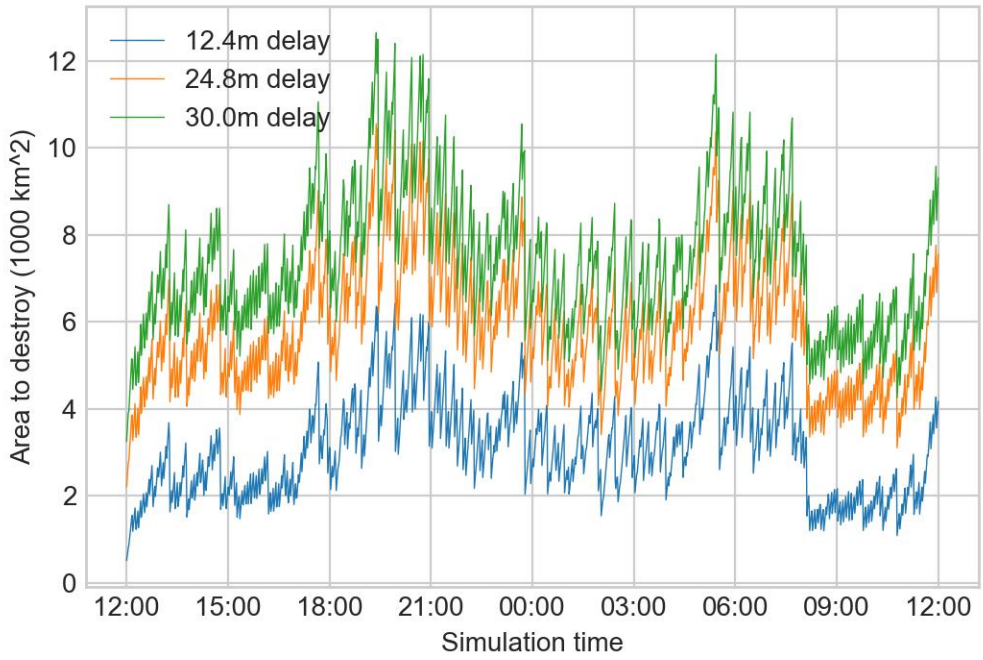
Second, our model also enables us to dig carefully into the mechanics of each run. At low alert, our model indicates China lacks a secure second-strike capability either way – no mated TELs survive US counterforce efforts. Examining our simulation output carefully, this result obtains because most Chinese TELs are not on deterrent patrol under low alert, and are in fact associated with specific bases to which they return and at which they can be destroyed. Additionally, most TELs are de-mated under low alert. On net, this makes human analysts even unaided by AI enough to locate any roaming TELs a significant fraction of the time under our low alert assumptions.

Nonetheless, even at low alert, the effects of AI versus human processing are tangible in the practical details of counterforce. The below figures show base case counterforce area under low alert as a function of simulation time, with and without AI. In particular, the AI-enabled graph is much more jagged because AI intelligence-processing increases the sensitivity of US counterforce efforts to moment-by-moment variations in available intelligence sources. Weather, time of day, and random chance affect the availability of US observations of Chinese TEL location; since exploitation of this knowledge is gated by human analyst attention without AI, however, the variation occurs over a longer timescale as the US intelligence community's humans produce each piece of finished intelligence. This patterns also obtains when observing undetected roaming time.

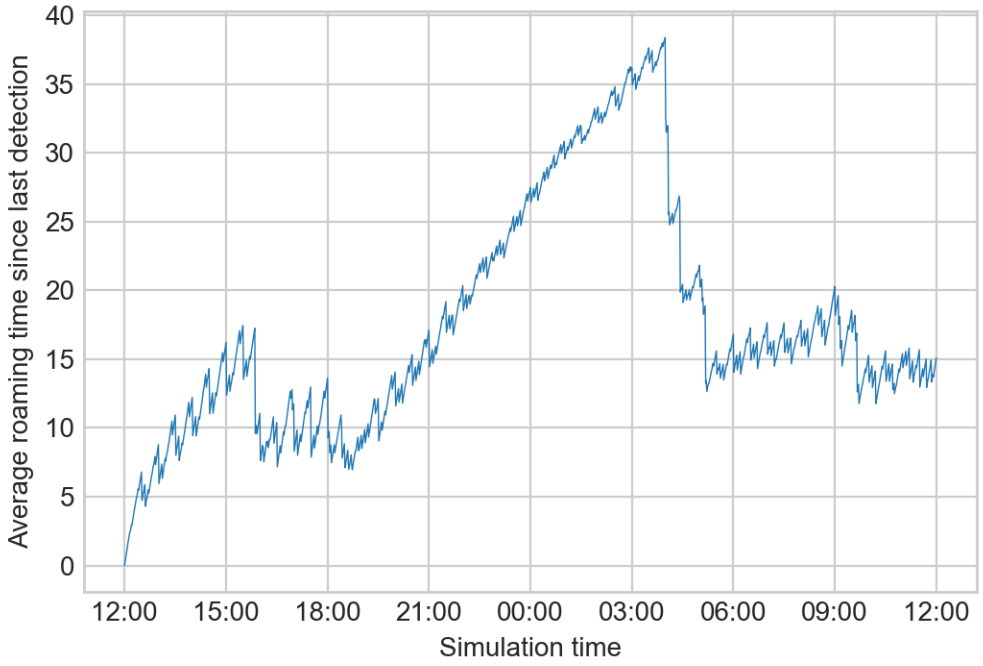
**Counterforce Area (Base Case, Low Alert, No AI)**



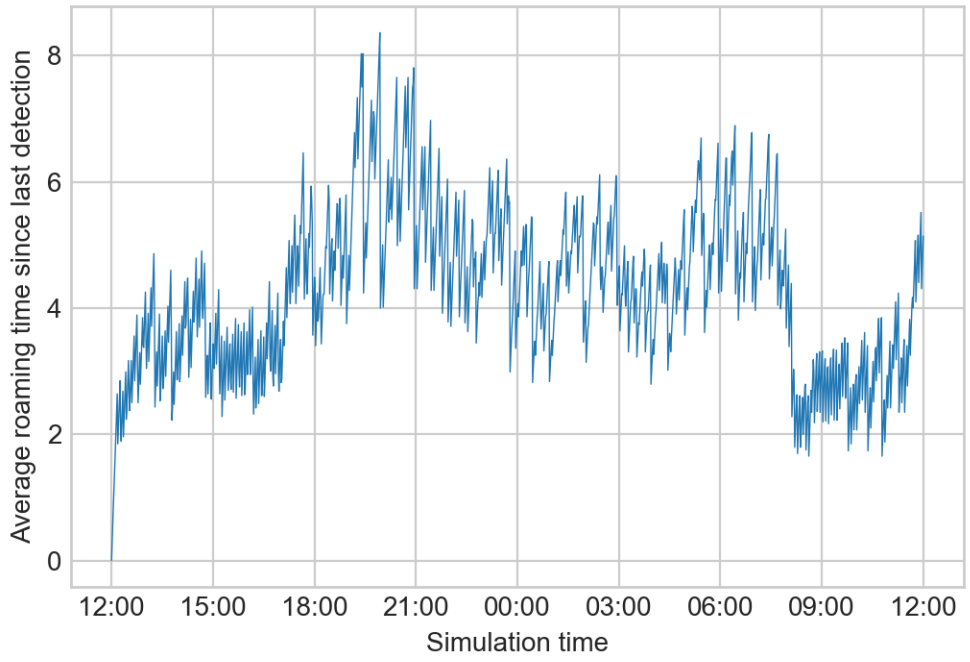

**Counterforce Area (Base Case, Low Alert, AI)**



**Undetected Roaming (Base Case, Low Alert, No AI)**

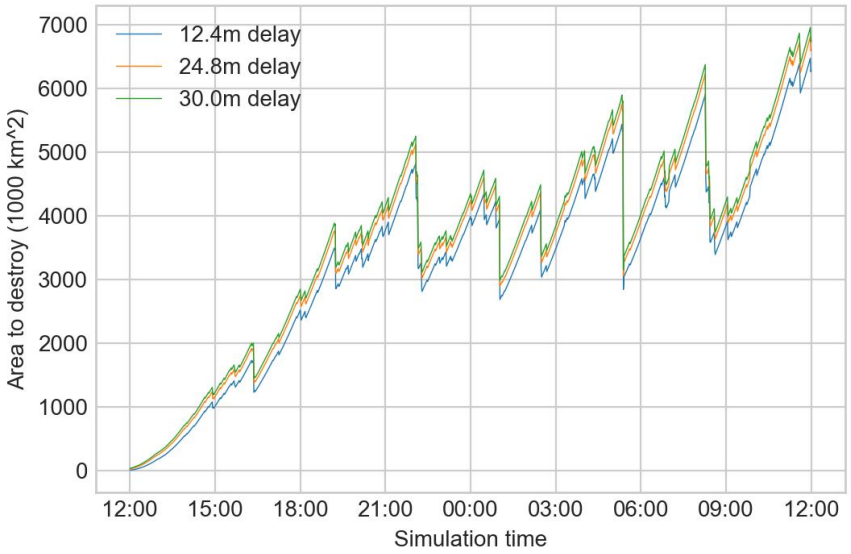



**Undetected Roaming (Base Case, Low Alert, AI)**

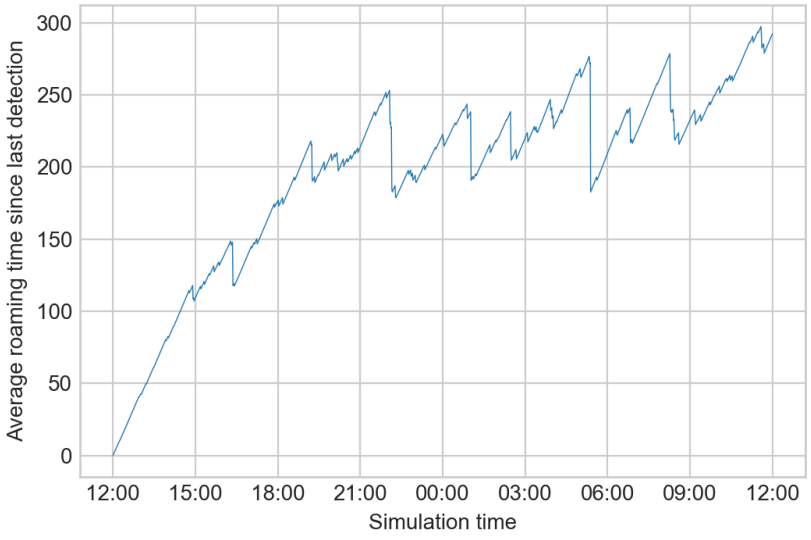


Moving to medium alert, we find that without AI, humans are no longer able to keep pace with Chinese TEL movements. Consequently, the counterforce area and average undetected roaming time continuously increase, and the probability of retaliation quickly converges to 1. Throughout the simulation duration, under these conditions about 20-30 mated TELs would survive US counterforce efforts, with some minute-to-minute variation based on the model's stochastic components.

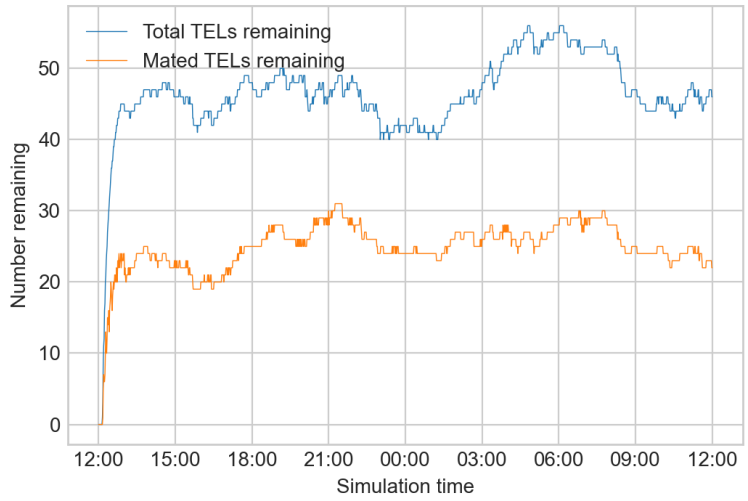
**Counterforce Area (Base Case, Medium Alert, No AI)**



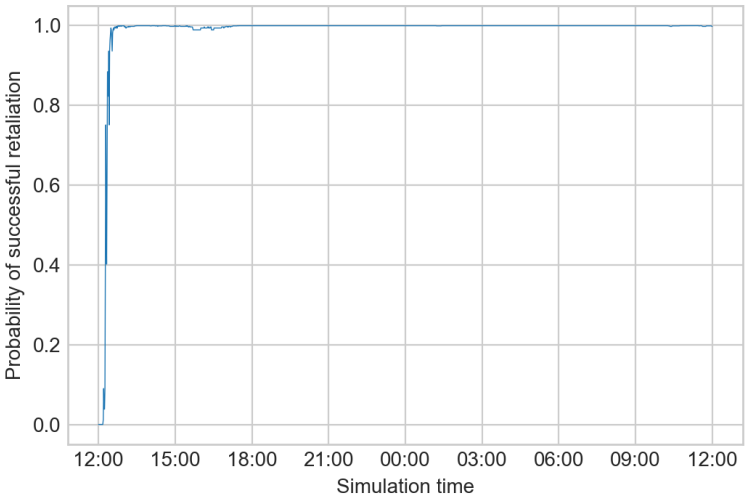

**Undetected Roaming (Base Case, Medium Alert, No AI)**



**Surviving TELs (Base Case, Medium Alert, No AI)**



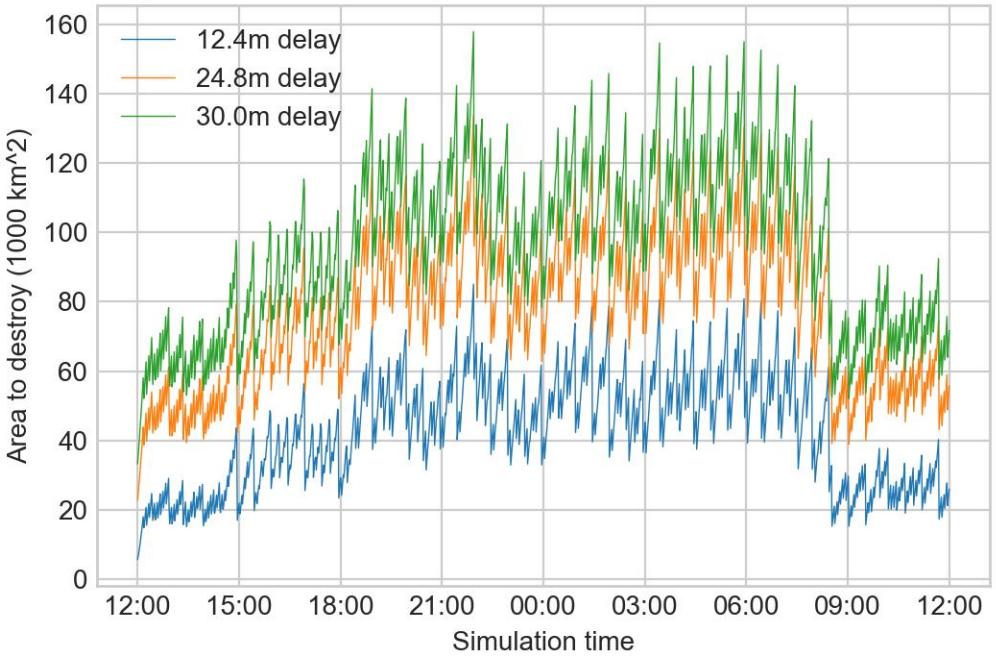

**Retaliation Probability (Base Case, Medium Alert, No AI)**



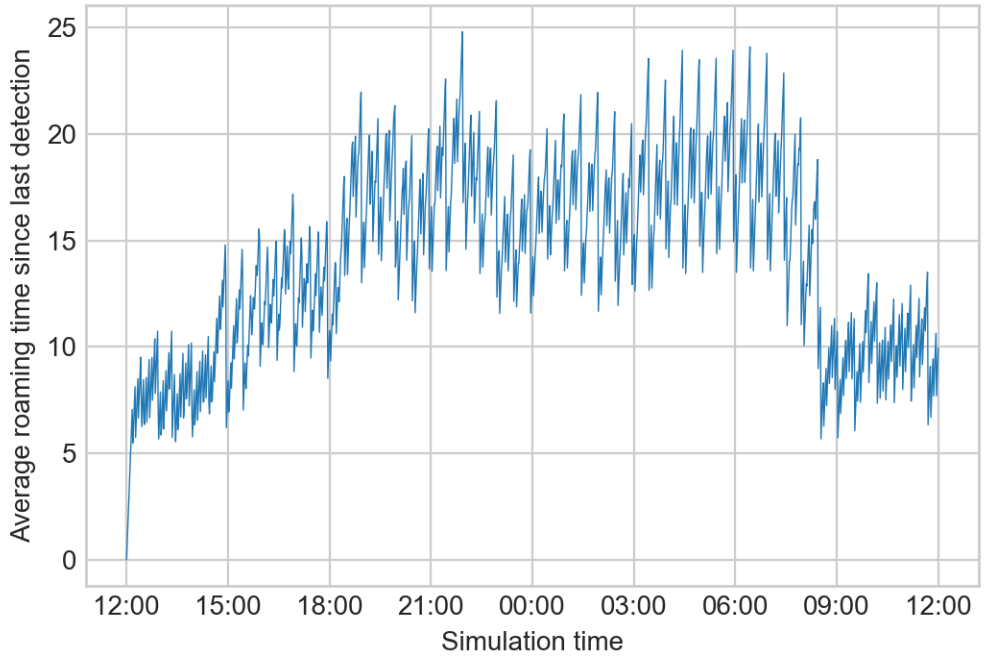

## (2) Base Case, Medium Alert, With AI

In contrast, with AI, the US still experiences occasional windows of opportunity where splendid first-strike against the Chinese nuclear arsenal is possible. Since these windows are brief and somewhat unpredictable, essentially depending on stochastic fluctuations in detecting and tracking the last marginal TEL, however, the United States would have to move toward an extremely ready posture, waiting for optimal striking time over a designated 24-hour period; such preparations would potentially be visible to China, prompting countermeasures to increase TEL survivability (such as by moving to high alert). To hold open these windows, the United States could also likely cue more targeted intelligence assets to engage in riskier, costlier, and/or lower-duration ISR missions (e.g., stealthy UAV missions, HUMINT for TELs in bad-weather areas, tailored SIGINT for particular TEL units and personnel) during likely future opportunities for first-strike (e.g., good-weather days).

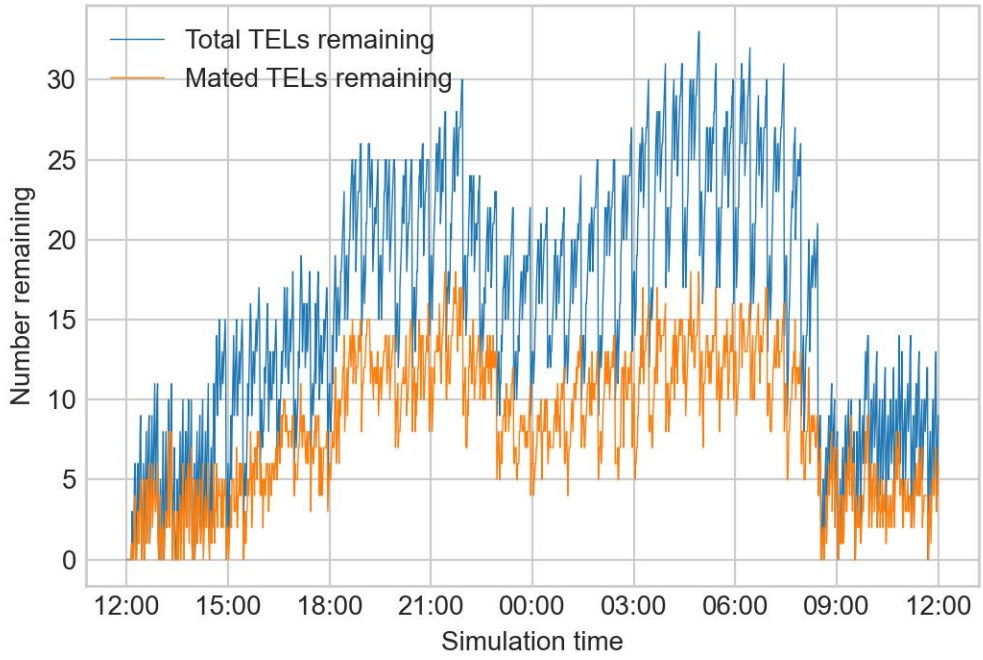
**Counterforce Area (Base Case, Medium Alert, AI)**



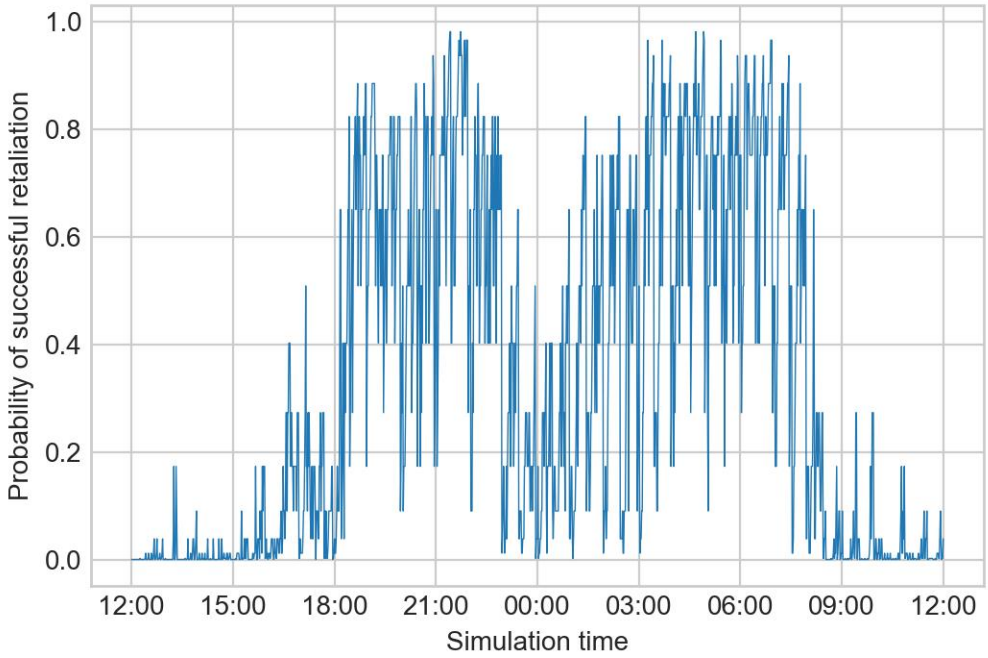

**Undetected Roaming (Base Case, Medium Alert, AI)**



**Surviving TELs (Base Case, Medium Alert, AI)**



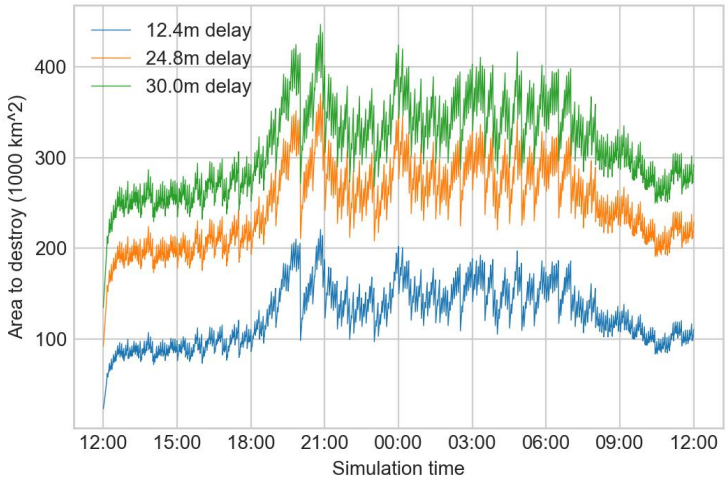

**Retaliation Probability (Base Case, Medium Alert, AI)**



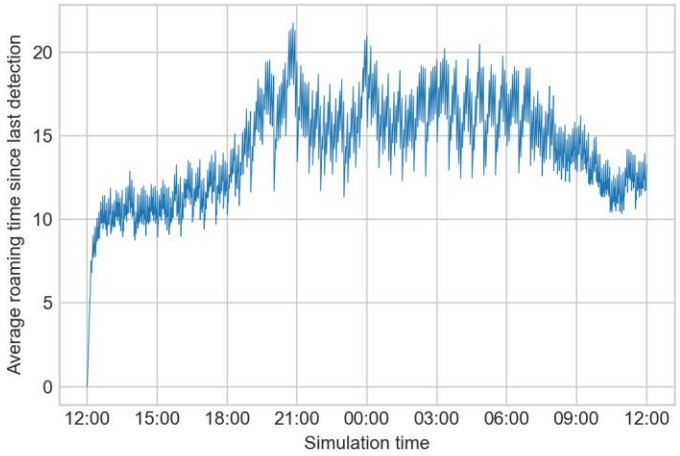
### (3) Base Case, High Alert, With and Without AI

In the high alert case, where we also give China credit for a number of countermeasures (e.g., 100% emissions control), we find that US counterforce efforts are unworkable either with or without AI. Digging into the raw results, we find this is primarily because of the combination of high TEL speed and irreducible latency – even with AI, there is an absolute minimum time needed to collect and analyze data; following that, there is some minimum flight time for US nuclear weapons, even depressed-trajectory SLBMs located within East Asia. Consequently, even when stochastic factors fluctuating over the course of simulation time give the United States relatively up-to-date information about TEL locations, a splendid first-strike remains out of reach. To combat this, the United States would have to pre-position strike assets just off China's coast, or acquire a larger nuclear arsenal and thereby become able to hold a larger area at risk with barrages.

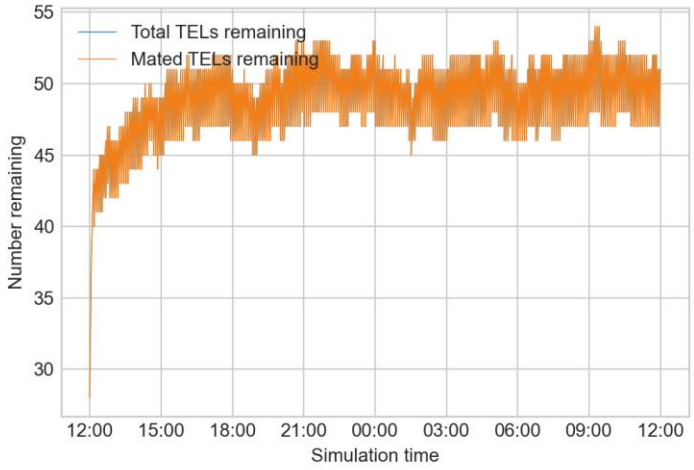
**Counterforce Area (Base Case, High Alert, AI)**



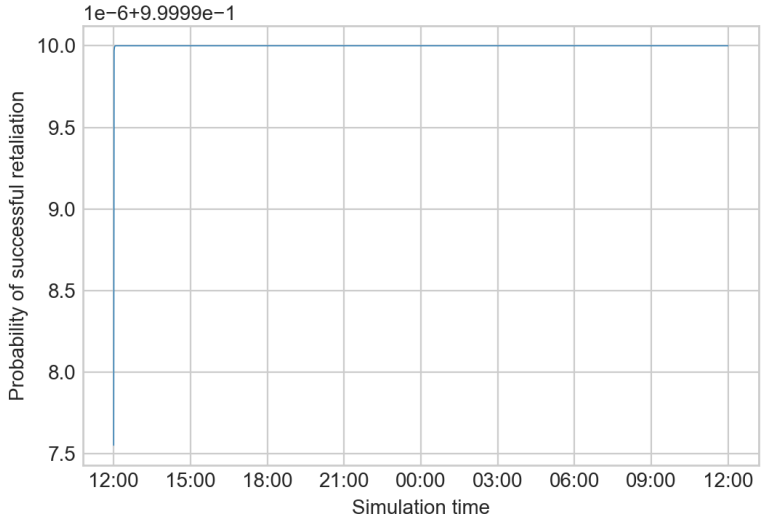

**Undetected Roaming (Base Case, High Alert, AI)**



**Surviving TELs (Base Case, High Alert, AI)**




**Retaliation Probability (Base Case, High Alert, AI)**

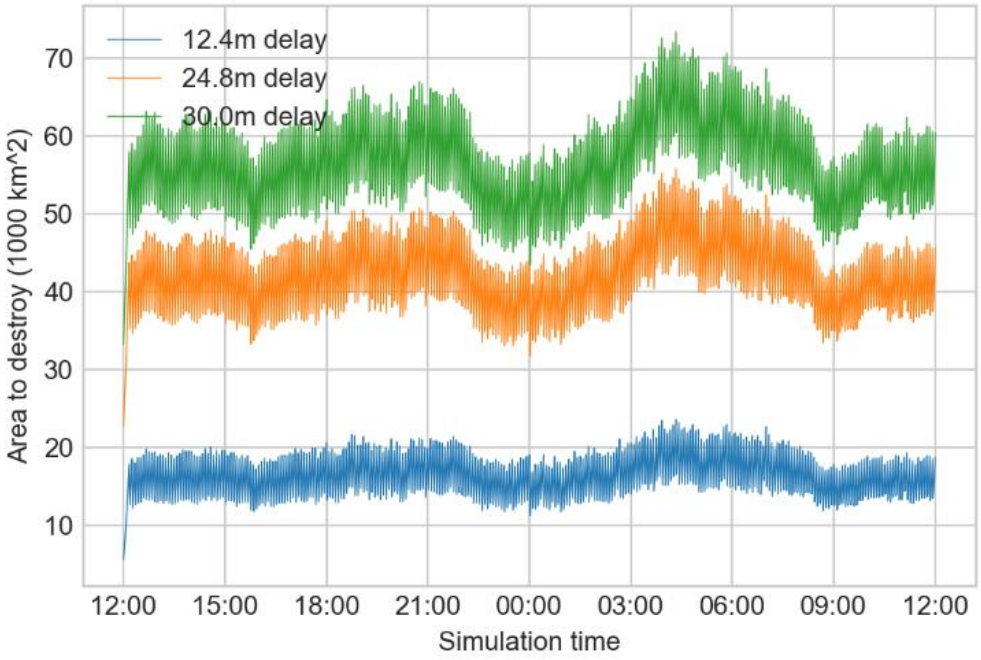



## (4) Excursions

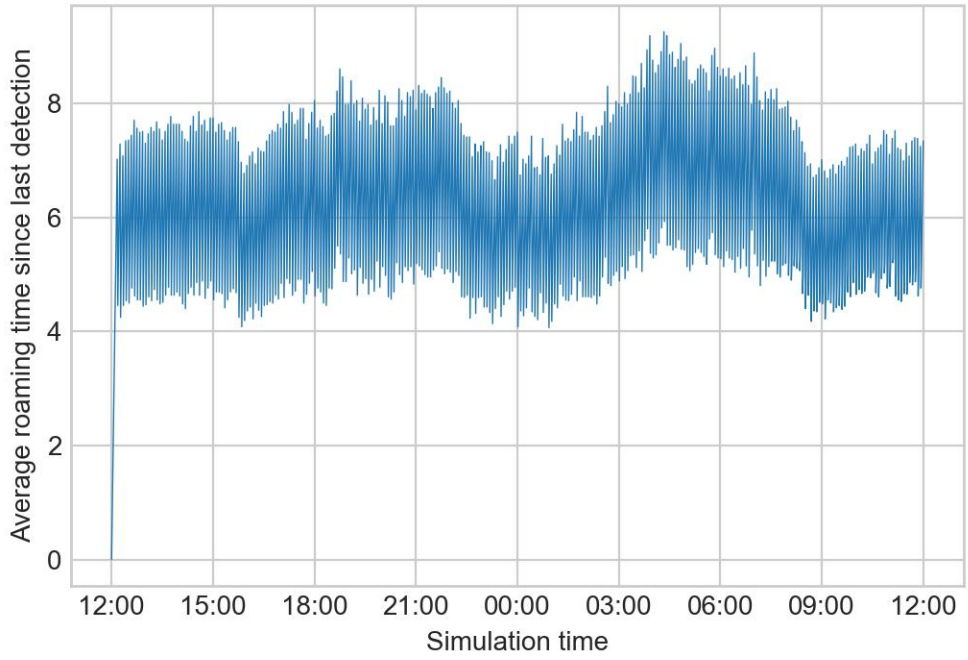
We also ran various excursions. First, the price of maintaining SAR constellations continues to fall, both due to cheaper launches and miniaturization. We assess the impact of a nearly saturated SAR constellation for the United States, where time between five-minute passes is reduced to five-minutes. Thus, in this excursion, the United States has day/night, all-weather SAR/GMTI coverage of most of China, topography permitting. Successful counterforce at low alert is trivial, as even the base case without AI was successful.

At medium alert, however, we find a significant change – instead of a mostly survivable Chinese deterrent with windows of opportunity for a US first-strike, the inverse obtains: US first-strike is mostly possible, outside a few stochastically generated intervals.

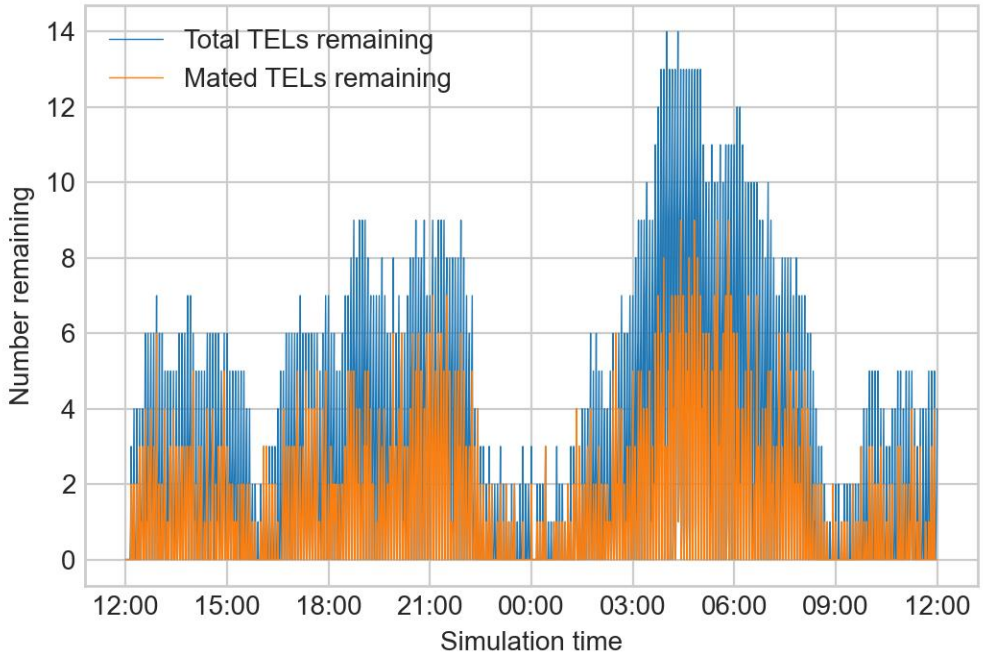
**Counterforce Area (SAR Excursion, Medium Alert)**



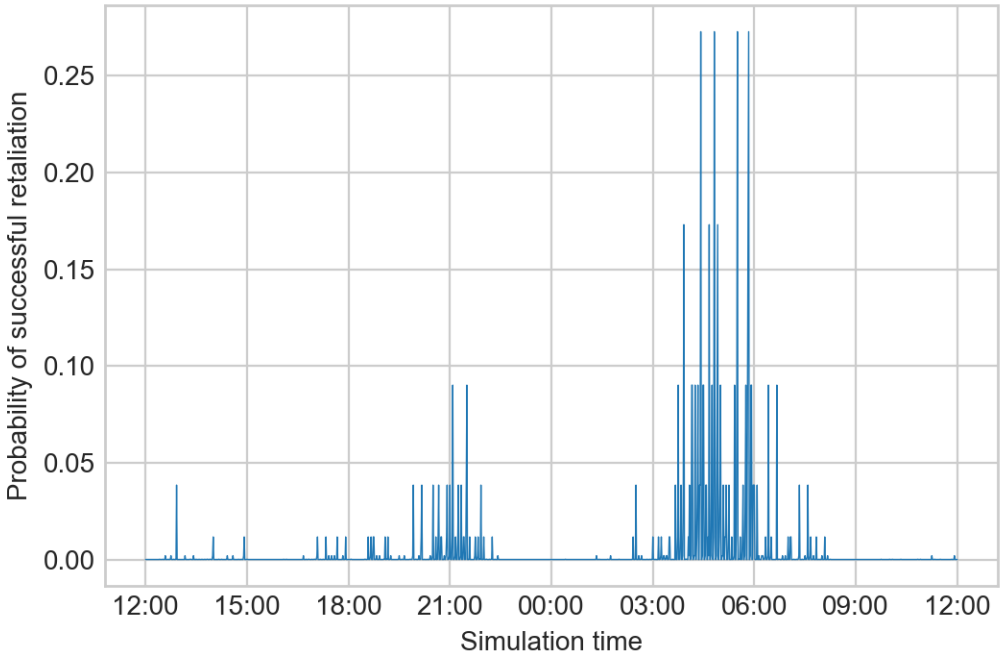

**Undetected Roaming (SAR Excursion, Medium Alert)**



**Surviving TELs (SAR Excursion, Medium Alert)**

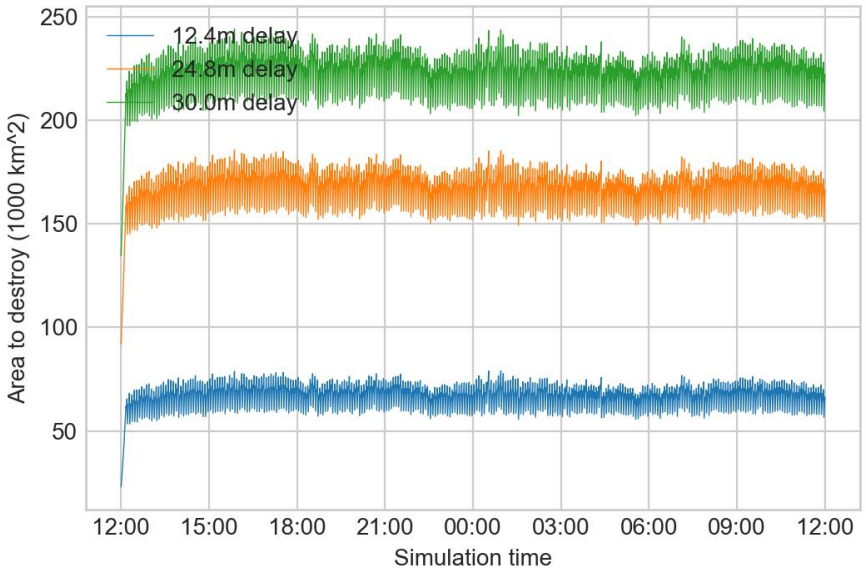



**Retaliation Probability (SAR Excursion, Medium Alert)**

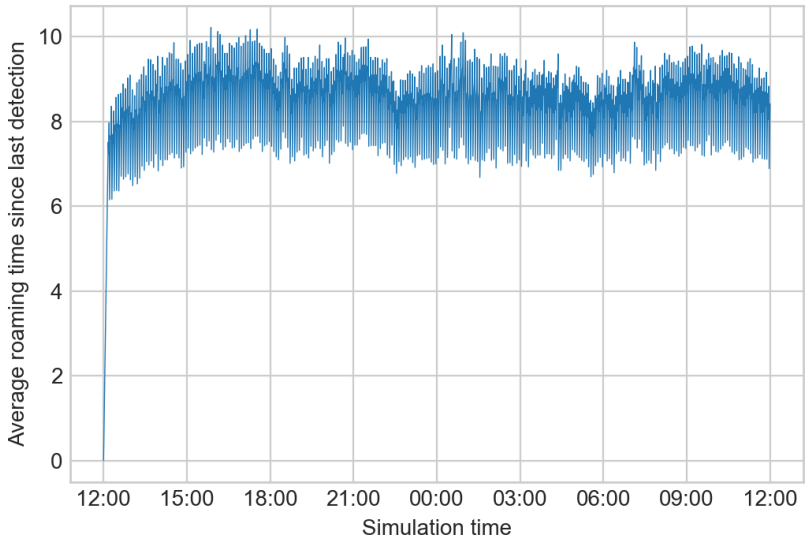


However, at high alert, as with the base case, we find again that US counterforce are still unworkable, due to high TEL speed and irreducible latency. This is a notable finding, as it suggests that even with AI, the progressive proliferation of miniaturized satellites will not itself threaten Chinese second-strike. Instead, to threaten a China which adopted the measures we bundle with high alert status, the US would have to itself change its force structure or take other costly actions.

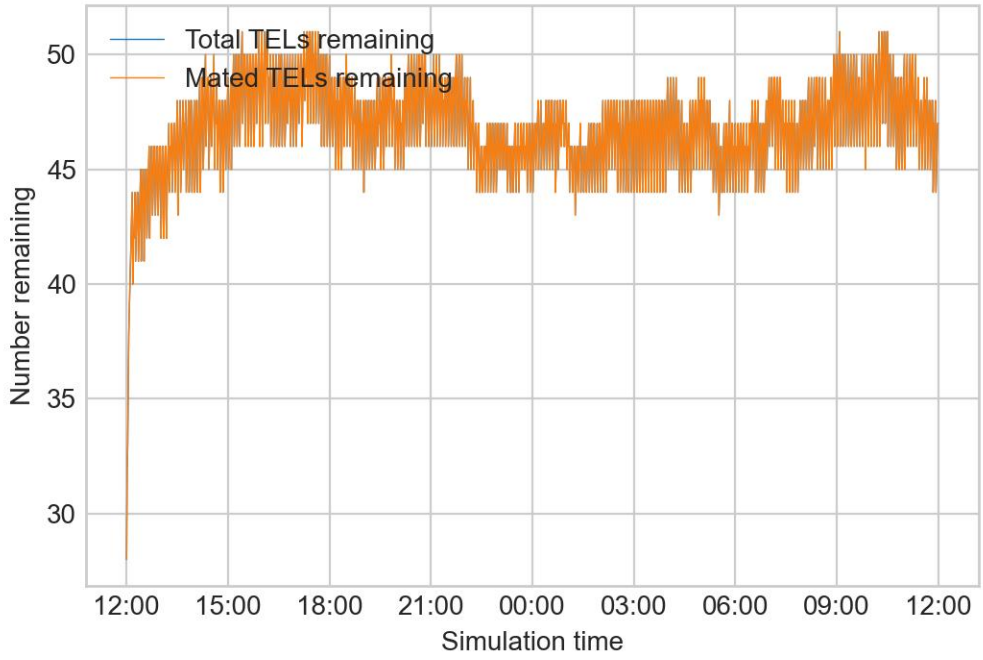
**Counterforce Area (SAR Excursion, High Alert)**



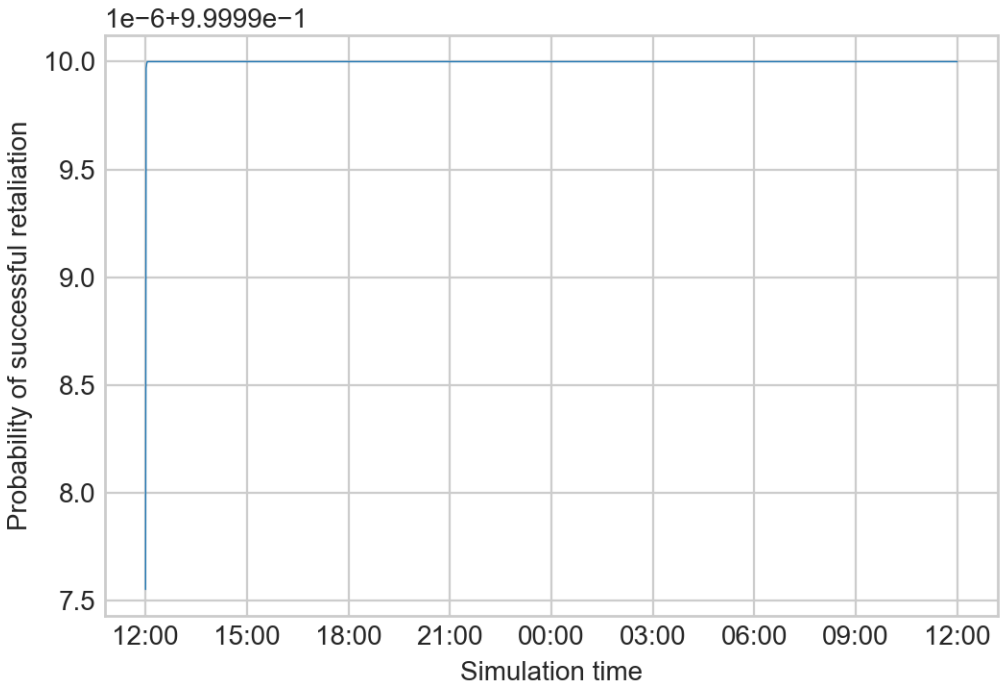

**Undetected Roaming (SAR Excursion, High Alert)**



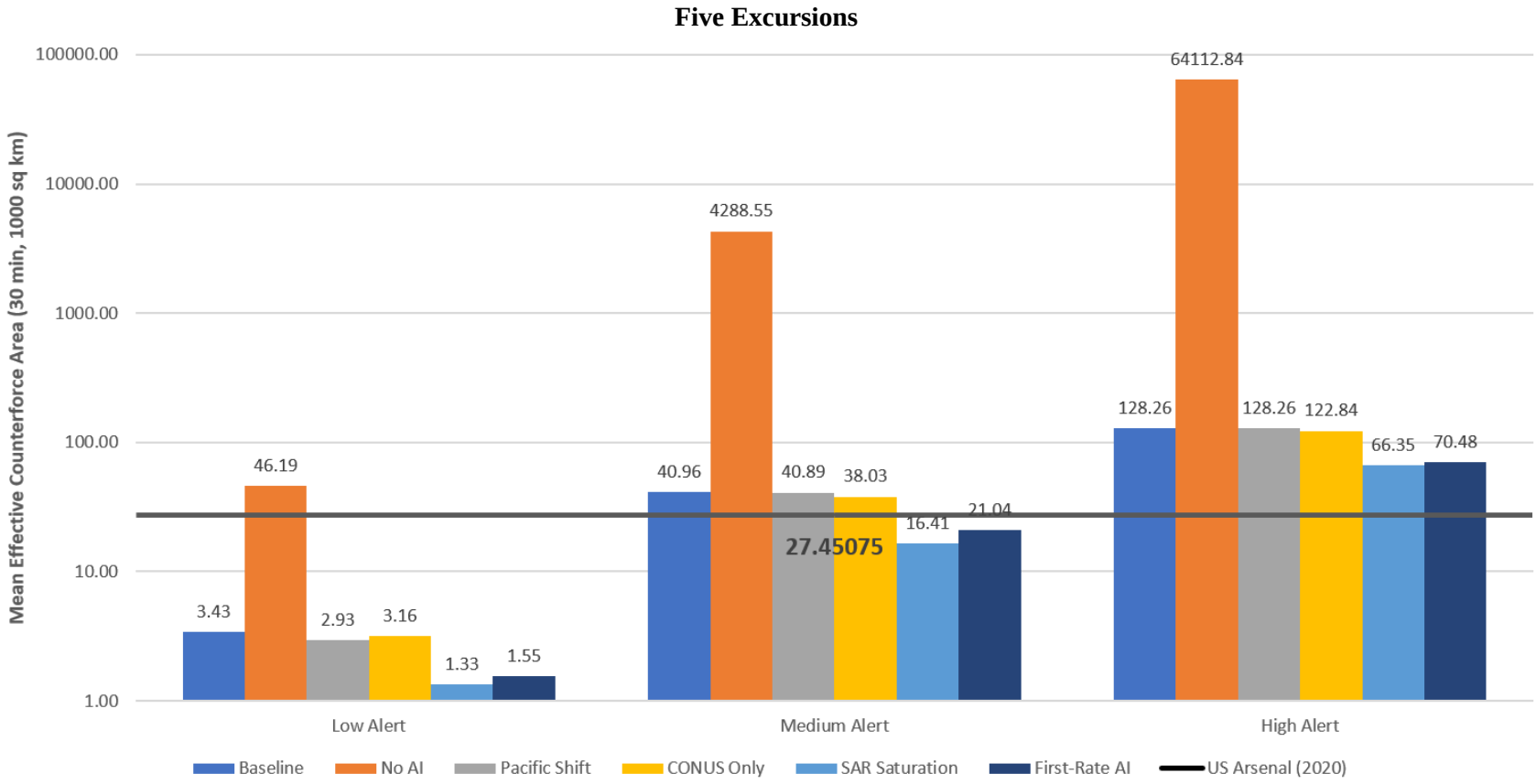
**Surviving TELs (SAR Excursion, High Alert)**




**Retaliation Probability (SAR Excursion, High Alert)**



Besides increased SAR satellites, we also asked if shifting more submarines to the Pacific, adopting a doctrine of being willing to only protect CONUS, or developing first-rate AI able to classify intelligence five times as fast as our point estimate would decisively change our results. The impact of developing first-rate AI was roughly comparable to launching more SAR satellites, and also permitted counterforce against China at medium alert; no excursions examined, however, enabled counterforce at high alert.



## Countermeasures

If our model successfully identifies that AI-enhanced counterforce will pose significant risks to China's arsenal, how might China respond? Nuclear weapons states have employed three main strategies to ensure the survivability of their arsenals: redundancy, hardening, and concealment.[388] Given the large size of the US nuclear arsenal relative to China's, redundancy and hardening have not been favored relative to concealment. Insofar as AI systems increasingly complicate concealment, however, China may feel compelled to consider countermeasures in all three categories (a-c). I discuss each in turn, as well as the possibility of China shifting to a launch-on-warning posture (d).

## (1) Concealment

As a strategy for survivability, concealment intuitively consists of limiting enemy knowledge about the locations of key aspects of one's nuclear weapons system. Presently, mobile platforms enjoy "concealment by default" when on deterrent patrol, because the total area they could be is very large relative to their size. In theory, TELs could be on any road in China, and submarines could be anywhere in the ocean. As our model illustrates, however, AI-enabled intelligence-processing may pose severe difficulties for this form of concealment, as rapidly and accurately ascertaining platform locations will become increasingly easy. To restore the efficacy of concealment, China could undertake several possible countermeasures. Conceptually, these countermeasures can essentially be reduced to two types: confronting an increase in US capability to track its mobile platforms, China could seek to restore its previous level of survivability either by degrading that capability or by increasing its own elusiveness.

### (a) Degrading US Capabilities

China could seek to degrade the newfound US ability to defeat concealment through mobility by attacking at several different points along the kill chain. Below, I discuss counter-AI and counter-space options.

#### Counter-AI

First, most obviously, China could seek to defeat US AI directly. As this dissertation's first paper discusses at length, deep learning systems are vulnerable to both training-time attacks, which seek to teach the AI system the wrong lessons (likely through breaches of cybersecurity), and test-time attacks, which seek to exploit imprecision in AI systems once trained.

On the former, access to US systems could enable Chinese hackers to teach image classifiers to see TELs as normal trucks, for example, or otherwise misclassify key intelligence.[389] While the ability to disable US systems through cyber-vulnerabilities is not new to AI, what is arguably notable is the scale of disruption – since AI does most of the work in our model of making the Chinese arsenal vulnerable, having one's AI system hacked means the US loses most of its counterforce capability. Thus, moving a greater percentage of the total "labor" of one's system to AI effectively increases the total vulnerability of one's military capabilities to cyber-intrusions.

On the latter, test-time attacks after training are arguably less likely without also having compromised cybersecurity anyway, as while adversarial examples are often discussed in the existing literature (e.g., small perturbations can cause image classifiers to see pandas as gibbons, for example), such attacks require fairly precise knowledge of the particular weights the AI system has learned to assign to different kinds of evidence, anyhow; that is, since the details of US counterforce-involved AI systems would almost certainly be highly classified, exploiting the possibility of adversarial examples would likely require some degree of

---

[388] Lieber and Press, "The New Era of Counterforce," 16-8.

[389] Andrew Lohn, "Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity," *CSET*, 2020, https://cset.georgetown.edu/research/hacking-ai/.

compromise in security in the first place, at which point less exotic attacks are also already available. Thus, we assess that while counter-AI efforts could help China restore nuclear survivability, they would depend strongly on Chinese ability to penetrate US nuclear systems with its capabilities in cyberspace.

Finally, as a third possibility, China could also seek to attack the interface between human beings and their AI systems. For example, perhaps exploiting AI's increasing ability to generate human-quality writing at scale, Chinese propagandists could promulgate narratives about the dangers of reliance on fickle, mysterious, and easy-to-deceive AI. If US elites cannot justify the expense involved in integrating AI systems into counterforce efforts to their publics, or if US military brass do not themselves trust their AI, then high-stakes counterforce efforts which rely on those systems will be unlikely. Of course, this attack modality is necessarily more speculative.

**Counter-Space**

Further, since the AI system modeled above relies on continuous streams of data from EO and SAR satellites, China could utilize counter-space capabilities to disable wartime tracking of its mobile platforms. Notably, China possesses several counterspace capabilities it could use against US satellites, including direct-ascent kinetic anti-satellite weapons (ASATs), as well as radio-frequency jammers and other directed-energy weapons.[390] Could blinding US EO and SAR satellites ensure TEL survivability? We assess the prospects as mixed.

First, pre-deploying counter-space assets would likely play into US efforts in intelligence preparation of the battlefield. Even in the base scenario, dazzling imaging systems requires co-locating dazzlers with the target being protected, however; the same holds for radio-frequency jamming. Consequently, if not also widely deployed across areas without TELs, the fact of dazzlers or jammers could itself help reveal TEL locations, or at least prompt further US scrutiny.[391]

Second, China could resort to ASATs. Even if China knew the United States would likely retaliate, since the latter generally relies far more on orbital infrastructure compared to China, China may feel that mutual denial of space would be advantageous. Consequently, Stephen Biddle and Ivan Oelrich assess that "in a high-stakes confrontation with China, [the United States] cannot assume its satellites will survive."[392] However, this would of course represent a destructive kinetic attack against US assets, and hence be a fairly escalatory move; it would also not guard against a bolt-from-the-blue attack of the sort modeled by Heginbotham et al. (2015) and Wu (2020), in addition to this paper, as baseline assessments, even if such an attack is politically unlikely on the part of the United States.

Further, as satellites continue to miniaturize, the resilience of US space assets is likely to sharply increase over the near to medium term, as the progressive miniaturization of satellite technology has enabled a super-linear growth in constellations of tiny satellites for all purposes.[393] In sum, strategic competition between

[390] Heginbotham et al., "The U.S.-China Military Scorecard," 245-57, 227-43.

[391] Ibid., 256.

[392] Stephen Biddle and Ivan Oelrich, "Future Warfare in the Western Pacific: Chinese Antiaccess/Area Denial, US AirSea Battle, and Command of the Commons in East Asia," *International Security* 41.1 (2016), 46.

[393] Christopher Mims, "The Tiny Satellites That Will Connect Cows, Cars and Shipping Containers to the Internet," *Wall Street Journal*, January 9, 2021, https://www.wsj.com/articles/the-tiny-satellites-that-will-connect-cows-cars-and-shipping-containers-to-the-internet-11610168400; Roberto Di Pietro, "The Coming Satellite Revolution: New Business Opportunities, Scenarios, and Threats," *Modern Diplomacy*, 2021, https://moderndiplomacy.eu/2021/05/14/the-coming-satellite-revolution-new-business-opportunities-scenarios-and-threats/.

the United States and China will also involve space, which in turn will affect US ability to leverage AI systems which depend on data pipelines involving satellites. Settling which way that balance will turn is outside the scope of this paper.

### (b) Increasing Chinese Elusiveness

Instead of degrading US AI or space capabilities, China could also seek to instead respond to an AI-augmented US ability to defeat concealment through mobility by seeking to amplify the elusiveness of its mobile platforms.

**"Denial through Silence"**

First, China could seek to defeat US AI-assisted counterforce by denying US AI systems the data needed to function, a strategy we might term "denial through silence." Since deep learning generates military power by substituting for, and sometimes exceeding, human beings at various data-intensive tasks, severely restricting what data an adversary can feed their AI system can constrain any AI-derived advantages. This denial could occur during both training and inference.

In training, deep learning systems need many examples of some given phenomena before becoming reliably able to perform identification. In fact, since deep learning systems are much more data-hungry than human beings, denial through silence can be effective against AI systems where the same tactic would have failed against humans. In the nuclear domain, this could take the form of, for example, excluding new TEL models from military parades, so as to minimize available data about their appearance, weight, and other specifications, making it harder to train US AI systems seeking to sort through intelligence to identify them. Indeed, this aspect of AI-related competition is likely to broadly incentivize increased peacetime secrecy about military capabilities – states broadly face tradeoffs, in considering whether to reveal the possession of some given capability, between generating deterrent and other value by signaling and the fear that enemies may develop potent countermeasures.[394] If a growing ensemble of AI systems depend for their efficacy on the obtainability of a key minimum quantity of data about adversary military systems, then incentives to conceal such systems so as to avoid erosion of their usefulness will increase.

In inference, in order for AI systems to observe and locate mobile platforms, they need to emit data streams on which the AI systems can do their work. China could reduce the surface area of such streams in several ways, as incorporated into the simulation of the different stipulated alert levels in our model. Of course, this principle is not unique to AI – in the undersea balance, for example, a constant struggle obtained during the Cold War between Soviet efforts to quiet submarines and American efforts to again track them, often by identifying increasingly exotic and obscure acoustic signals.[395] Although this would be somewhat costly in terms of ensuring reliable command-and-control, units could be given more authority to vary their paths in times of crisis, reducing the degree to which US intelligence preparation of the battlefield efforts could usefully extrapolate from data about past behavior. Further, China could instruct its TELs (and, in theory, its nuclear-armed submarines) to practice emissions control, communicating only minimally during deterrent patrols.

**"Denial through Noise"**

Obversely, China could seek to hide any true signals revealing TEL location amidst noise. First, as we model above, China could deploy large numbers of decoys, which at perfect indistinguishability would be

---

[394] Brendan Rittenhouse Green and Austin Long, "Conceal or Reveal? Managing Clandestine Military Capabilities in Peacetime Competition," *International Security* 44.3 (2019/2020), 48-83.

[395] Cote, "The Third Battle"; Long and Green, "Stalking the Secure Second Strike."

equivalent to building up its arsenal size, as any US counterforce effort would then also be obligated to strike all decoys. To approach indistinguishability, however, would require a fairly significant effort on the part of the PLA, as TELs at present have various highly distinctive pattern-of-life tells which set them apart from civilian traffic, such as the presence of support vehicles (to say nothing of living at specific bases).

Second, China could also repeat this exercise for other avenues of intelligence, such as by transmitting large numbers of falsified orders about TEL movement back and forth on channels known to be monitored by the US intelligence community. It is difficult to assess the efficacy of such efforts now, but given the increased quality of AI-generated text and video, China could in theory begin to itself use AI to create noise, perhaps generating fake activity many times over compared to any signs of real TEL movements. Which side prevailed would then depend on relative expertise at leveraging their respective military AI systems.

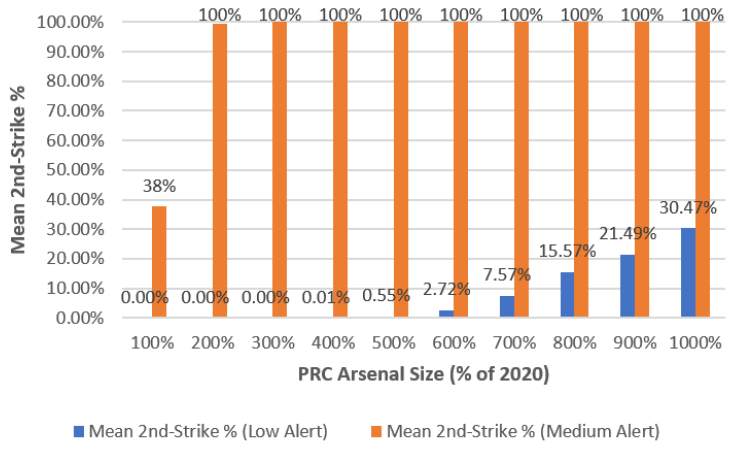
### (2) Redundancy

Redundancy refers, simply, to increasing the number of targets which must be disabled by any adversary to obviate one's retaliatory capability. I discuss a simple increase of China's arsenal size, as well as reliance on other deliver systems from the other two legs of the triad.

#### (a) Arsenal Size

First, most obviously, China could simply increase the size of its nuclear arsenal. The DoD diagnoses exactly such an effort to be underway, forecasting that China will at least double its stock of warheads over the next decade.[396] Strikingly, however, our model finds that doubling the Chinese arsenal only matters under medium alert, increasing the mean probability of successful second-strike from 38% to 100%. Under low alert, the Chinese arsenal must be sextupled before breaking a 1% probability of successful second-strike; under high alert assumptions, Chinese second-strike capabilities are already secure.

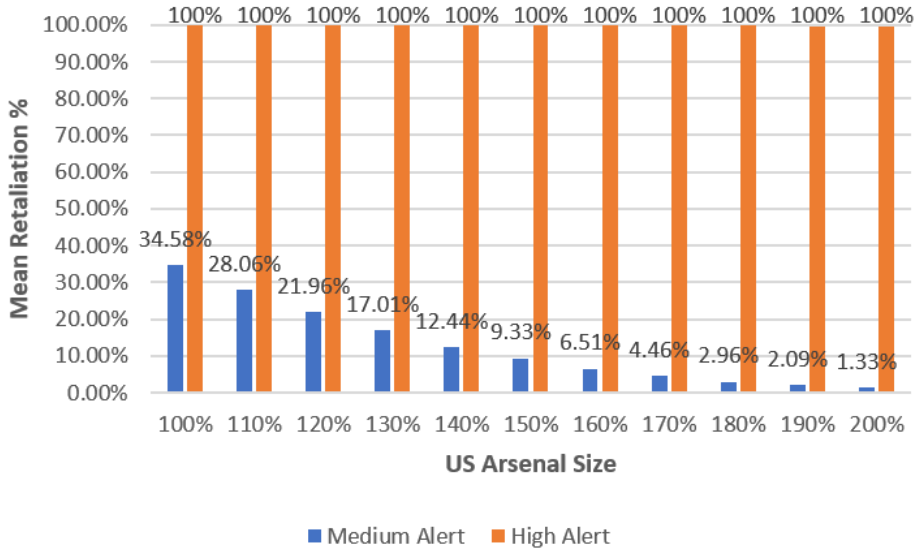
**Successful 2nd-Strike Probability and PRC Arsenal Buildup**




Conversely, considering high alert, even doubling the US nuclear arsenal does not drop the chance of successful Chinese second-strike below 100%; at medium alert, a decrease is palpable, but a full 150% capability is required before the chance of retaliation drops below 10%. Of course, even a further 50% buildup of the already-large US nuclear arsenal is extremely unlikely.

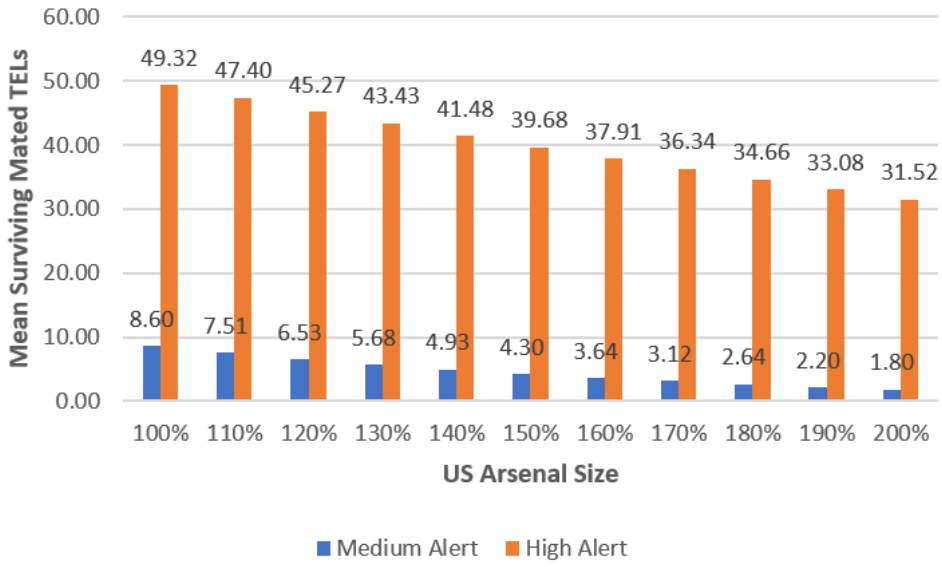
[396] Idrees Ali and Phil Stewart, "Pentagon concerned by China's nuclear ambitions, expects warheads to double," *Reuters*, September 1, 2020, https://www.reuters.com/article/us-usa-china-military-nuclear/pentagon-concerned-by-chinas-nuclear-ambitions-expects-warheads-to-double-idUSKBN25S5MB.

**Higher PRC Alert Levels Are Robust to US Arsenal Buildups – Second-Strike Probability**



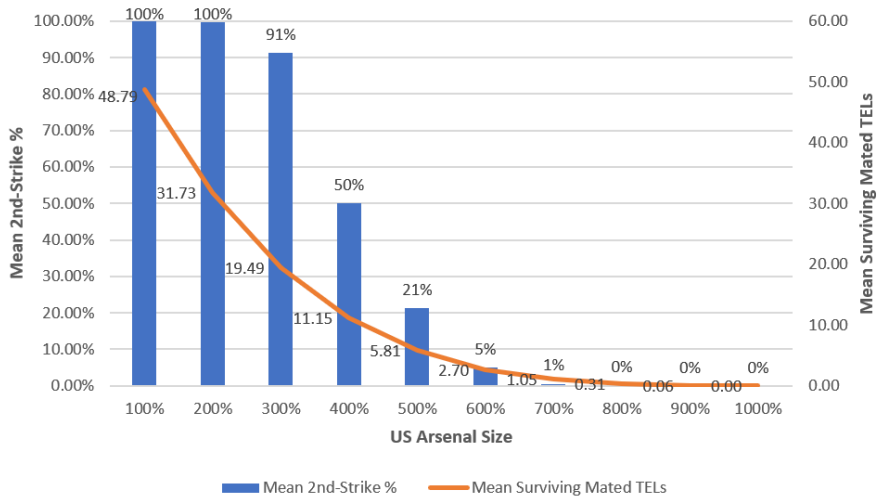

This dynamic can also be seen in examining the number of mated TELs which, on average, survive a US counterforce effort – here, on high alert, fully doubling the US nuclear arsenal only purchases the destruction of about 18 TELs; TELs under medium alert are comparably costly.

**Higher PRC Alert Levels Are Robust to US Arsenal Buildups – Mated TELs**



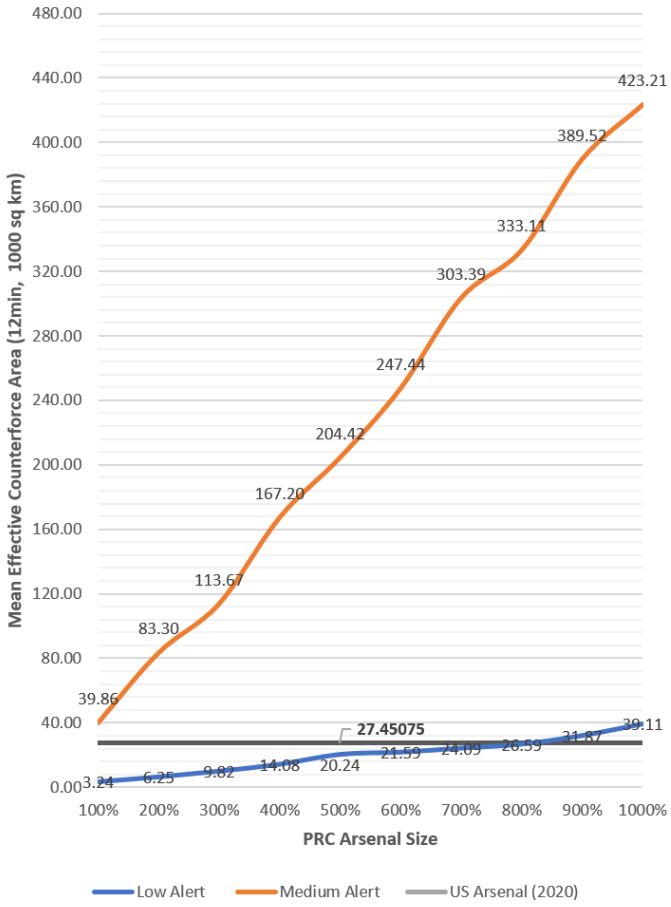

Asking our simulation for the break point, we find that the US arsenal would have to be sextupled – an increase of over 10,000 deployed warheads – to confidently threaten China's second-strike capability at high alert.

**US Arsenal Size Against Mean 2nd-Strike % and Mated Surviving TELs**




In short, alert levels decisively dominate arsenal size in our model, in either direction. To illustrate why this is the case, we graph the effective counterforce area produced by PRC arsenal increases of up to 1000% the original baseline case, at both medium and low alert:

**Effective Counterforce Area – Low and Medium Alert**



Here, the solid black line represents the mean area a US nuclear counterforce effort can destroy TELs across (assuming a 2:1 targeting doctrine, airburst, and 5 psi), averaging across stochastic effects and normalizing to a 12-minute flight time to account for weapon latency. The blue and orange lines, equivalently, represent the total mean target area presented by the Chinese nuclear arsenal under low and medium alert, respectively; when these lines exceed the solid black line, this means that some TELs would survive the average run of the simulation.[397] Intuitively, what is occurring is that each additional TEL under a higher alert level is worth a much larger amount than the same TEL under a lower alert level – the difference in slope is so large that the blue line would have to be continued very far off the graph to achieve what a single doubling does for the effective counterforce area of the PRC arsenal, under medium alert.

**(b) Delivery Systems**

I discuss submarines and bombers in turn. First, as mentioned above, the consensus of both American and Chinese analysts is that the United States possess a strong lead against Chinese submarines, having previously benchmarked against Soviet submarines quieter than current-generation Chinese assets.[398] Wu similarly dismisses Chinese submarines out of hand, noting that US advantages in the undersea domain are "likely to persist for a long time," owing to both a large technological lead and the likely ease of monitoring submarine transits around the first island chain.[399]

Second, while China's H-6 bomber could theoretically represent an air leg finishing its nuclear triad within the decade, this would face several barriers to representing increased survivability against US AI-enabled counterforce efforts. First, most obviously, airfields, airstrips, and hangars all represent fixed targets, easily located now via amateur use of Google Maps, let alone AI-enabled processing of all US intelligence streams. Since one or two nuclear weapons would almost certainly suffice to destroy each airbase, this could represent only marginal increases to the effective counterforce area presented to US counterforce efforts. Second, several factors would complicate airborne alert. During a period of the Cold War, the United States ran 24/7 airborne alert missions using nuclear-armed B-52s, so that a portion of the US nuclear deterrent would still survive and be available for retaliation even if every square inch of American earth were irradiated by Soviet missiles. These efforts were discontinued after eight years, however, because of the high frequency of accidents, including one which caused the radioactive contamination of Denmark and another which resulted in the temporary loss of several weapons to the ocean.[400] Given the PRC's relatively small arsenal size, combined with the Politburo's revealed preferences about the always/never dilemma up to the present, this would likely seem an especially costly way to purchase additional survivability.

Further, even assuming the prospect of AI-enabled counterforce put the fear of God, so to speak, into the CCP, an effort to credibly strike CONUS with nuclear-armed bombers would face severe range difficulties.

---

[397] Note that this is not equivalent to successful second-strike, in part because the US possesses missile defenses which are also implemented in the simulation, and also because taking average effects in effect treats US decision-makers as randomly initiating a counterforce attack at some arbitrary point in the day. In practice, of course, decision-makers would only agree to counterforce attacks with high chances of success, and so would selectively pick from more favorable parts of the stochastic distribution (e.g., weather, chance SIGINT intercepts, low civilian truck traffic).

[398] Zhao, "Tides of Change"; Wu, "Survivability of China's Sea-Based Nuclear Forces"; Cote, "Assessing the Undersea Balance Between the U.S. and China."

[399] Wu, "Living with Uncertainty," 106-7. There are countermeasures. For example, improvements in the range of Chinese SLBMs could theoretically enable a bastion strategy inside the first island chain. We hope to model the effect of AI on the undersea balance in future work.

[400] Rebecca Grant, "The Perils of Chrome Dome," *Air Force Magazine*, 2011, https://www.airforcemag.com/article/0811dome/.

Assuming H-6s equipped with China's air-launched ballistic missile (ALBM), the CH-AS-X-13, the bomber would need to fly about 7,000 km to be able to strike the western coast of the United States. Since the H-6 has a claimed maximum combat radius of 3,500 km, this would require aerial refueling even assuming the pilot would ditch into the ocean after firing. This would be, to say the least, a fairly awkward way of carrying out nuclear retaliation - for one, for example, given the H-6's maximum speed of about 1000 km/h, the bomber would need to take on a heroic 7-hour journey to range California, giving the United States a very long window in which to intercept. Of course, China could in theory solve this problem by both acquiring an airbase closer to the United States, and then basing nuclear-armed bombers there, but this would increase the difficulties involved in instituting some form of 24/7 airborne alert (otherwise, the United States could likely strike that forward-positioned airbase with conventional capabilities at the start of any kinetic conflict).[401]

Finally, restoring survivability through redundancy would require vast non-warhead investments. In particular, since any US nuclear counterforce effort can always target the weakest link in the overall Chinese nuclear weapons system, it does no good to build 1,000 warheads if Chinese command-and-control can still be easily destroyed. Thus, a redundancy strategy would also require installing additional communications networks, command and control systems, and all other aspects of the Chinese nuclear weapons system other than the warheads and launch vehicles themselves.

### (3) Hardening

Third, China could increase the difficulty of destroying individual TELs. While the model already assumes TELs hardened to require 5 psi of overpressure to destroy, and in targeting solutions assigns US weapons sufficient to blanket each targeted TEL with 5 psi twice, faster TELs could function as better weapon sponges. That is, examining our model, we found that a key contributor to large effective counterforce areas at medium and high alert was, simply, TEL speed – although AI cut out the vast majority of the information-processing latency between "find" or "fix" and "finish," the weapons used to "finish" China's road-mobiles still need to travel between 12 and 30 minutes to reach their target, by which time they must strike a radius proportional to TEL driving speed. Even with perfect AI-assisted ISR, gains in counterforce ability are still bounded by the geography of one's strike assets. The Pentagon's Scud hunt report noted, "Planners cannot target things whose existence is unknown to them"; after AI, perhaps we should note, "Even when their existence is known, planners cannot efficiently target things moving at high speeds that are far away."[402]

While TELs cannot drive at their maximum speed on all roads at all times, as this would surely risk accident and incur unacceptable maintenance costs, there is the simple brute fact that TELs driving at high speeds, no matter how effective the AI-empowered information-processing, can represent very large barrage targets simply by virtue of drawing a large circle. Consequently, modifying TELs to drive faster and more sustainably may be a simple way for China to recover some portion of survivability, either alone or alongside other packages of actions coded as part of our alert level schema. Under perfect transparency, it is worth noting that mobility represents a form of hardening, not concealment.

[401] "Military and Security Developments Involving the People's Republic of China 2020: Annual Report to Congress," DoD, 2020, https://media.defense.gov/2020/Sep/01/2002488689/-1/-1/1/2020-DOD-CHINA-MILITARY-POWER-REPORT-FINAL.PDF.

[402] Watts and Keaney, "Effects and Effectiveness," 330.

### (4) Posture

Finally, China could also recover survivability by switching to a launch-on-warning posture, as the DoD's 2020 report to Congress forecasts.[403] At best, such a doctrinal change could enable a high degree of protection against counterforce efforts, albeit at the cost of leaving some nuclear systems on high alert. Nonetheless, some reasons point against the efficacy of such a measure.

First, an early warning system might not substantially increase the acceptable launch window for a retaliatory Chinese strike. After all, while the United States uses various forward-deployed assets as part of its early warning system, a Chinese equivalent would likely, at least at first, be much more primitive and rely on assets based on Chinese soil.[404] Further, the travel times for US nuclear weapons are inherently very short. For US submarines located in the Pacific, depressed-trajectory SLBM strikes would take about 12 minutes on average to reach their targets, and positioning close to the Chinese coast could theoretically reduce this travel time to a single-digit number. Thus, relative to existing doctrine, switching to a launch-on-warning posture would only drag forward in time a few minutes the allowable window within which Chinese retaliation could occur. In fact, Wu and Li assert that even Chinese solid-fuel missiles take "tens of minutes" or "half an hour" to become ready to launch; while fixing these vulnerabilities would doubtless be part of adopting a launch-on-warning posture, sufficiently low weapon latencies could once again enable some form of damage limitation capability.[405]

Second, this doctrinal shift to launch-on-warning would require massive across-the-board changes to Chinese command-and-control, training procedures, and default readiness conditions. That is, given the above extremely abbreviated notice, to make a Chinese nuclear attack against the US homeland possible with just a few minutes' warning would require day-and-night changes across the entirety of the Chinese nuclear weapons system. Obviously, Chinese TELs would need to be mated ahead of time, and any confirmatory procedures presently in place to prevent accidental launch and confirm that orders were flowing from China's national command authority would have to be minimized. Since Chinese elites currently prefer the "never" branch of the always/never dilemma, culturally dislike pre-delegation, and generally favor negative control, shifting to a launch-on-warning posture would represent a large organizational/cultural shift of the sort likely requiring adaptation on the scale of years, rather than months.

Finally, in the framework of long-term strategic competition, China's assured retaliation posture has enabled it to enjoy a relatively tame nuclear balance with the United States, as well as the resource savings from abstaining from a large build-up a la US-Soviet nuclear competition. Keeping nuclear forces on continuous alert to make launch-on-warning possible would increase accident risk, generate large and continuous organizational demands on Chinese military and civilian elites, and would require a burst of sustained spending on various aspects of the Chinese nuclear weapons system.[406]

---

[403] "Military and Security Developments Involving the People's Republic of China," *DoD*, 2020, https://media.defense.gov/2020/Sep/01/2002488689/-1/-1/1/2020-DOD-CHINA-MILITARY-POWER-REPORT-FINAL.PDF#page=110.

[404] Cunningham and Fravel, "Assuring Assured Retaliation," 31.

[405] Wu, "Appendix," 4-5; Li, 11.

[406] Wu, "Living with Uncertainty," 115-7.

## Conclusion

China is large, and mobile launchers are small. Since 2007, China has leveraged this basic fact to great effect, militating against large US advantages derived from arsenal size, overall military spending, and most of all the long experience of nuclear competition with the Soviet Union. As our simulation shows, however, advances in modern AI are eroding this previously easy source of nuclear survivability. In our model, AI provides an order-of-magnitude or more reduction in counterforce requirements. Mobility is no longer a guarantee of obscurity, as deep learning classifiers obviate the need for human analysts to process image, text, audio, electronic, and other intelligence flows. If TELs are present in the massive, continuously generated dataset incoming from satellites, drones, intercepts, and other sources, the AI state-of-the-art has reached the point where a classifier can likely be trained to find them at superhuman speed.

In the simplest sense, the US arsenal is large relative to all states but Russia. Consequently, if the US can exchange its warheads for those of a target state at any reasonable, reliable rate, it has a strong chance of possessing a nuclear counterforce ability against that state. Every other aspect of both the US and other states' nuclear weapons systems adjusts what that exchange rate may be, but at the heart of the matter, the US arsenal is large, the Chinese arsenal is small, and China cannot realistically deploy thousands of additional weapons. AI's decisive effects flow from this core dynamic: without concealment, adjusting that exchange rate will require much more effort than China has previously put into nuclear competition with the United States. Redundancy and hardening run up against US arsenal size. Our model allows us to conclude that based on open-source information about US intelligence assets and AI capabilities, road-mobile missiles are indeed eminently detectable and trackable under various plausible conditions.

## Nuclear Survivability After AI

What, then, can China do? Two main possibilities remain: negating the new AI capability itself, through either digital or physical means, and shifting to a launch-on-warning posture. Many ways in which China could conceivably negate the new AI capability, however, are also themselves escalatory. Digitally speaking, in the field, fooling AI systems requires understanding them, which means compromising the cybersecurity of US nuclear assets. Data poisoning attacks definitionally entail direct modification of the AI's training process, after all, while the construction of adversarial examples is difficult without knowledge of the trained AI model's weights. Thus, an intensified need to defeat adversary AI systems may worsen what Ben Buchanan terms the "cybersecurity dilemma," where states feel compelled to penetrate each other's systems in cyberspace for defensive reasons, leading to the mutual perception of hostility.[407]

Physically speaking, destroying US satellites *en masse* would be an act of war. Pre-delegating launch authority to TELs would undermine the credibility of China's no-first-use pledge, and higher TEL speeds, radio silence, and constant deterrent patrols would all raise the risk of accidents. Shifting to launch-on-warning would require significant investment in, and retooling of, China's overall nuclear weapons systems, heralding a new era of nuclear competition between the United States and China that the world has thus far been able to avoid. Consequently, the possibility of AI-enabled US nuclear counterforce raises issues policymakers will have to carefully consider and navigate. Notably, while worries about accident risk related to Chinese delegation of nuclear command and control to AI are oft-discussed, none of these specific measures involve AI.[408] AI's ripple effects are not limited to the domain of AI.

---

[407] Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (UK: Oxford University Press, 2016).

[408] Saalman, "Fear of false negatives." During the Cold War, as is oft cited as precedent, both the United States and the USSR used earlier forms of AI to pursue greater automation in nuclear command and control, including for logistical planning, missile targeting, and guidance. Most infamously, the Soviet Union deployed a Dead Hand system which would autonomously initiate retaliation if a nuclear weapon had struck the Soviet Union and the

Further, if AI enables successful US counterforce against China, its effects likely also matter for imaginable US counterforce efforts against Iran or North Korea. According to reporting by Reuters, for example, the US has exactly begun applying AI to finding and tracking North Korean mobile missiles.[409] Of course, while China's arsenal size is small relative to the United States, North Korea and Iran are even worse off. Elsewhere, no hard barrier precludes other states facing nuclear adversaries with relatively small arsenals from, in theory, adopting AI-enabled counterforce strategies. India has recently begun exploring the possibility of counterforce against Pakistan.[410] South Korea may come to seek a conventional counterforce effort against North Korea, with the steady acquisition of UAVs, EO satellites, and a miniaturized SAR constellation beginning in 2025.[411]

Generally speaking, if AI means the death of mobility as a relatively cheap way of achieving arsenal survivability, and if even conventional counterforce efforts become increasingly plausible, we may see the reemergence of global nuclear competition among many dyads, as previously survivable arsenals become less so. For dyads with highly asymmetric arsenal sizes, the warhead-poor state may have to adopt some of the countermeasures we explore here for China, such as defeating the AI directly or switching to a more hair-trigger launch posture, since building up a larger inventory may be impractical and even invite preemption. For dyads where weapons are fewer, a race to build more weapons may be reignited.

Simultaneously, the increasing efficacy of AI-assisted counterforce may also deter new would-be nuclear states from attempting to acquire weapons, due to the increased difficulty of protecting nascent arsenals from preventive attacks. If AI renders even fairly mature arsenals vulnerable to counterforce, crossing the threshold of acquiring a single operational weapon may no longer be sufficient to guarantee against homeland invasion.[412] Rather, new nuclear states may be forced to ford a longer period of vulnerability.

Future work should investigate the possibility of AI-enabled counterforce against states other than China, both by the United States and by regional actors. In addition, we do not investigate in this paper all AI applications which could assist US counterforce efforts. First, this dissertation's third essay argues that AI hardware advantages mean the US will better leverage any hypothetical drone swarming based on advanced deep learning algorithms; some analysts believe drone swarms would significantly ease conventional damage limitation efforts, as their mass deployment both undersea against submarines and in air against strategic bombers would not be constrained by a nation's supply of appropriately skilled human beings.[413]

---

communications link to the Soviet General Staff had been destroyed. See D. E. Hoffman, *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy* (Anchor Books: New York, 2009).

[409] Stewart, "Deep in the Pentagon, a secret AI program to find hidden nuclear missiles."

[410] Clary and Narang, "India's Counterforce Temptations."

[411] Ian Bowers and Henrik Stalhane Hiim, "Conventional Counterforce Dilemmas: South Korea's Deterrence Strategy and Stability on the Korean Peninsula," *International Security* 45.3 (2020/21), 23-4, https://doi.org/10.1162/isec_a_00399.

[412] Alexandre Debs and Nuno P. Monteiro, *Nuclear Politics: The Strategic Causes of Proliferation* (UK: Cambridge University Press, 2017).

[413] Zachary Kallenborn, "Are Drone Swarms Weapons of Mass Destruction?" Center for Strategic Deterrence Studies, 2020, https://media.defense.gov/2020/Jun/29/2002331131/-1/-1/0/60DRONESWARMS-MONOGRAPH.PDF; Zachary Kallenborn and Philipp C. Bleek, "Drones of Mass Destruction: Drone Swarms and the Future of Nuclear, Chemical, and Biological Weapons," War on the Rocks, February 14, 2019; Zachary Kallenborn and Philipp C. Bleek, "Swarming destruction: drone swarms and chemical, biological, radiological, and nuclear weapons," The Nonproliferation Review 25.5 (2018): 523-43; Zachary Kallenborn, "Meet the future weapon of mass destruction, the

Second, the AI-powered intelligence-processing modeled in this paper for nuclear counterforce would be equally useful for defeating concealment efforts in any conventional counterforce campaign. While only nuclear weapons can supply large-area barrage attacks against TELs in motion, if Wu (2020) is correct in assessing that Chinese mobile missiles cannot launch if all pre-prepared firing locations are destroyed, then deep learning classifiers would likely be of immense help in locating all such locations by continuously combing through intelligence data to detect new construction activity.[414]

Finally, AI may eventually help augment US missile defense efforts. The key technical difficulty with missile defense is decoys. In the atmosphere's near-vacuum, light and heavy objects move similarly, and sophisticated decoys are coated to heat, cool, and reflect identically to real warheads. Since incoming missiles travel at many times the speed of sound, targeting and launch are already automatic, but current algorithms rely on handcrafted features.[415] While data about what decoy measures or new missile types China might deploy are obviously highly secret, synthetic data or simply excellent intelligence could enable the US to bring AI to bear.[416]

Thus, more campaign analyses would shed further light on AI's nuclear effects. More generally, we hope the present effort advances the cause of simulation as a method in political science. By dynamically modeling the US intelligence process, TELs, and environmental factors, we were able to observe variation over 24-hour periods in US first-strike ability. For many kinds of military questions for which data is scarce, we suggest that simulation of this sort can begin to shed light on the issues at stake.

## Implications for Policymakers

Despite the inevitable simplifications in any simulation, we believe we amass significant evidence that China's nuclear deterrent, assuming realistic conditions, is fairly vulnerable to a US first-strike, both without – but especially with – AI. We find that AI has enormous effects on the US ability to locate China's road-mobile missiles, rendering thinkable the execution of the nuclear counterforce mission under mixtures of stochastic and strategic circumstances that would be one to several orders of magnitude too demanding otherwise. Since this effect derives from the speed, volume, and accuracy with which modern AI systems can process intelligence streams the United States already possesses, rather than from the deployment of some discrete novel capability, gradual adoption of such systems seems almost inevitable. Many capabilities which would further strengthen an AI-assisted nuclear counterforce campaign by the United States, such as the further development of robust SAR constellations, will likely proceed because they have many other non-nuclear uses.

This raises the challenge of managing strategic stability with China. China has thus far resisted trilateral nuclear arms control talks with the United States and Russia, a stance some observers have characterized as aggressive alongside the US intelligence community's recent assessment that China may double its nuclear arsenal over the next decade.[417] If China believes US AI-assisted counterforce would more than

drone swarm," Bulletin of the Atomic Scientists, April 5, 2021, https://thebulletin.org/2021/04/meet-the-future-weapon-of-mass-destruction-the-drone-swarm/.

[414] Wu, "Living with Uncertainty."

[415] Boulanin et al., "Artificial Intelligence, Strategic Stability and Nuclear Risk," 58.

[416] El-Darymli, "Automatic Target Recognition in Synthetic Aperture Radar Imagery: A State-of-the-Art Review," https://ieeexplore.ieee.org/abstract/document/7572958, especially part II; Ratches, J. A., "Review of current aided/automatic target acquisition technology for military target acquisition tasks."

[417] David Ignatius, "The wizards of Armageddon may be back," *The Washington Post*, 2021, https://www.washingtonpost.com/opinions/2021/05/06/wizards-armageddon-may-be-back/.

compensate for even a tripled arsenal, however, it may come to seek diplomatic negotiations as part of its next evolution in nuclear strategy. As rivalry between the United States and China is likely to continue intensifying over the near to medium term, both states must exercise caution to avoid dynamics which would raise the risk of a mutually catastrophic nuclear war. Especially as discussions about how to modernize the US nuclear arsenal continue in earnest, policymakers should privately review to what degree they intend to deliberately hold China's arsenal at risk, rather than stumbling into ever-more-effective counterforce capabilities by virtue of AI development undertaken for other purposes.

Further, our simulation constitutes significant evidence that AI is likely to impact nuclear balances more generally, including those outside the US-China dyad. Progressively cheaper access to imaging satellites, including in dyads such as India and Pakistan, means that access to data may not bottleneck the use of AI in any nascent counterforce efforts.[418] US policymakers considering the stabilizing management of regional rivalries should identify dyads where AI may induce first-strike instability, and coordinate with allies to minimize resulting inadvertent escalation risks.

[418] Clary and Narang, "India's Counterforce Temptations," 32.

# No Chips for the Drones of China: Why Hardware Will Bottleneck Chinese Military Artificial Intelligence

## Introduction

In Beltway discussions of artificial intelligence (AI), national security, and China, it has become impossible to avoid repeatedly encountering the following idea: AI may enable China to leapfrog US military power.[419]

But how? Naively, after all, the US vastly outspends China on essentially every relevant metric; concerns about Chinese home-field advantage are valid, but have nothing to do with AI. Nor do those raising the "AI leapfrog hypothesis" appear to uniformly believe that the gap between the US and China is small, whether owing to geography or other factors, and that therefore AI, simply being the latest in some long succession of new technologies relevant to military affairs, need only make up some small amount of ground. Rather, analysts seem to believe AI's effects could be fundamental.

Why? Although scholars' emphases differ, we can usefully crystallize the most prominent arguments underpinning the leapfrog hypothesis into the following two claims: (1) AI-enabled autonomous weapons ("AAWs") will possess decisive advantages over current-generation military assets, and (2) China will better leverage the existence of such weapons.[420] Consequently, the argument goes, the rise of such weapons in modern warfare could counterfactually enable Chinese victory in conflicts with the United States. In the words of Graham Allison, China may be "currently on a trajectory to overtake the United States in the

[419] Graham Allison, "Is China Beating America to AI Supremacy?," *The National Interest*, December 22, 2019, https://nationalinterest.org/feature/china-beating-america-ai-supremacy-106861; Robert O. Work and Greg Grant, "Beating the Americans at Their Own Game: An Offset Strategy with Chinese Characteristics," *CNAS*, June 6, 2019, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Report-Work-Offset-final-B.pdf?; Elsa B. Kania, "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power," *CNAS*, November 28, 2017, https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power; Gregory C. Allen, "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security," *CNAS*, February 6, 2019, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy; James Johnson, "The end of military-techno Pax Americana? Washington's strategic responses to Chinese AI-enabled military technology," *The Pacific Review* (2019), https://www.tandfonline.com/action/showCitFormats?doi=10.1080/09512748.2019.1676299, 7, 11; Michael Kanaan, *T Minus AI: Humanity's Countdown to Artificial Intelligence and the New Pursuit of Global Power* (Texas: BenBella Books, Inc., 2020), 192; Christian Brose, *The Kill Chain: Defending America in the Future of High-Tech Warfare* (New York: Hachette Books, 2020), xxv.

[420] For this paper, the phrase "AI-enabled autonomous weapons" ("AAWs") refers to weapons platforms using artificial intelligence to operate without human input into target selection after the moment of deployment. Conceptually, modern-day remotely piloted vehicles (RPVs) enter engagements without human pilots physically on board, but still require the cognitive labor of the remote pilot. AAWs are those theoretical future systems which would completely substitute AI for human beings as a source of that cognitive labor, enabling states to field arbitrary quantities of drones unbound by the scarcity of skilled human pilots. (To avoid confusion, unmanned aerial vehicles are defined as aerial vehicles where no human pilot is physically on board. Thus, both RPVs and AAWs would be considered UAVs.)

The phrase "AI-enabled autonomous weapons" seeks to avoid the tiringly extensive definitional disputes surrounding the more commonly used term, "lethal autonomous weapons systems" (LAWS). For example, it is contested as to whether landmines, which technically do not require human input for target selection after deployment, should count as LAWS. Separately, various weapons (e.g., missile defenses) have long since used varying degrees of "autonomy." Consequently, I use "AI-enabled autonomous weapons" to specifically indicate those weapons whose autonomy is AI-derived. For more, see Paul Scharre and Michael C. Horowitz, "An Introduction to Autonomy in Weapon Systems," *CNAS*, 2015, https://www.cnas.org/publications/reports/an-introduction-to-autonomy-in-weapon-systems.

decade ahead," assisted by "AI-empowered drone swarms" which could "make aircraft carriers … obsolete, all for one-thousandth of the cost."[421] Thus, the claim goes, we must outrace China to weaponize AI.

Notably, the first claim is not without its critics. Advocates argue AAWs gathered into drone swarms could, in theory, overwhelm human-piloted aircraft and surface vessels through sheer number, maneuverability, and/or decision-making speed, or at the least generate highly favorable cost exchanges.[422] Further, since losing AAWs would not carry nearly the same political cost as losing human lives, states without AAWs would risk losing contests of political endurance, whether in wartime itself or during peacetime episodes of coercion.[423] At the top end, some analysts argue advanced drones are best conceptualized as WMD.[424] Overall, on this view, states witnessing others' pursuit of AAWs will be obligated to acquire their own.[425]

---

[421] Allison, "Is China Beating America to AI Supremacy?".

[422] Ibid.; Christian Brose, "The New Revolution in Military Affairs: War's Sci-Fi Future," *Foreign Affairs*, May/June 2019, https://www.foreignaffairs.com/articles/2019-04-16/new-revolution-military-affairs; T. X. Hammes, "The Future of Warfare: Small, Many, Smart vs. Few & Exquisite?", *War on the Rocks*, July 16, 2014, https://warontherocks.com/2014/07/the-future-of-warfare-small-many-smart-vs-few-exquisite/; Paul Scharre, "Robotics on the Battlefield, Part I: Range, Persistence, and Daring," *CNAS*, 2014; ---, "Robotics on the Battlefield Part II: The Coming Swarm," *CNAS*, 2014; ---, "Robots at War and the Quality of Quantity," *War on the Rocks*, February 26, 2015, https://warontherocks.com/2015/02/robots-at-war-and-the-quality-of-quantity/; David Pinion, "The Navy and Marine Corps Need to Prepare for the Swarm of the Future," *War on the Rocks*, March 28, 2018, https://warontherocks.com/2018/03/the-navy-and-marine-corps-must-plan-for-the-swarm-of-the-future/; Andrew Ilachinski, "AI, Robots, and Swarms: Issues, Questions, and Recommended Studies," *CNA*, January 2017, https://www.cna.org/cna_files/pdf/DRM-2017-U-014796-Final.pdf; Robert O. Work and Shawn Brimley, "20YY: Preparing for War in the Robotic Age," *CNAS*, January 2014; Zhu Xiaoning, "Analysis of military application of UAV swarm technology," paper presented at the 3rd International Conference on Unmanned Systems (November 27-28, 2020, Harbin, China), https://ieeexplore.ieee.org/abstract/document/9274974; James Johnson, "Artificial Intelligence, Drone Swarming and Escalation Risks in Future Warfare," *The RUSI Journal* 165.2 (2020): 26-36; Norine MacDonald and George Howell, "Killing Me Softly: Competition in Artificial Intelligence and Unmanned Aerial Vehicles," *PRISM* 8.3 (2019), 102-27; Michael C. Horowitz, "Information-Age Economics and the Future of the East Asian Security Environment," in *The Nexus of Economics, Security, and International Relations in East Asia*, eds. Avery Goldstein and Edward D. Mansfield (US: Stanford University Press, 2012), 211; Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1.3 (2018), 47-8.

[423] Amy Zegart, "Cheap fights, credible threats: The future of armed drones and coercion," *Journal of Security Studies* 32.1 (2020): 6-46; Erik Gartzke, "Blood and robots: How remotely piloted vehicles and related technologies affect the politics of violence," *Journal of Strategic Studies* (2019); Erik Lin-Greenberg, "Remote Controlled Restraint: The Effect of Remote Warfighting Technology on Crisis Escalation" (PhD diss., Columbia University, 2019); ---, "Wargame of Drones: Remotely Piloted Aircraft and Crisis Escalation" (unpublished manuscript), 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3288988.

[424] Zachary Kallenborn, "Are Drone Swarms Weapons of Mass Destruction?" *Center for Strategic Deterrence Studies*, 2020, https://media.defense.gov/2020/Jun/29/2002331131/-1/-1/0/60DRONESWARMS-MONOGRAPH.PDF; Zachary Kallenborn and Philipp C. Bleek, "Drones of Mass Destruction: Drone Swarms and the Future of Nuclear, Chemical, and Biological Weapons," *War on the Rocks*, February 14, 2019; Zachary Kallenborn and Philipp C. Bleek, "Swarming destruction: drone swarms and chemical, biological, radiological, and nuclear weapons," *The Nonproliferation Review* 25.5 (2018): 523-43; Zachary Kallenborn, "Meet the future weapon of mass destruction, the drone swarm," *Bulletin of the Atomic Scientists*, April 5, 2021, https://thebulletin.org/2021/04/meet-the-future-weapon-of-mass-destruction-the-drone-swarm/.

[425] Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W. W. Norton and Company, 2018), 330; Brose, *The Kill Chain*, 115.

On the other hand, no current-generation drones have defeated human-piloted aircraft or bested modern air defenses on the battlefield; as a future technology, no empirical proof yet exists that AAWs will revolutionize warfare.[426] Since AAWs will require high-tech inputs, they may be sufficiently expensive that they fail to generate profitable cost exchanges when used as swarms.[427] Further, various countermeasures, such as air defense artillery, may be highly effective.[428] Overall, this debate seems difficult to resolve now, as discussions about how future and existing weapons systems will interact likely turn on hard-to-predict questions surrounding doctrines, capabilities, and countermeasures. The second claim, however, I argue, can be evaluated now.

To summarize, this paper's central claim is that the non-diffusion of AI hardware will prevent China from better leveraging, in the context of conventional conflict with the United States, any future AI-enabled autonomous weapons capability which emerges within at least the next ten years. Even stipulating that future AAWs will in fact defeat current human-controlled military assets, and even if China develops better doctrine and organizational practices for such weapons, the PLA will be fundamentally limited by lacking self-sufficient access to advanced AI chips.[429]

Importantly, this argument rests on several key assumptions. While I defend their likelihood below, it is worth making explicit the scope conditions without which my thesis might prove untrue:

- *US/ally relations.* I assume a minimum degree of US/ally coordination about export controls. Only a "minimum degree" is required because the United States itself controls several supply chokepoints, but the effectiveness of supply cutoff would be immensely reduced (to say nothing of other military problems) if all of East Asia collectively decided to bandwagon with China.

- *No radical technological breakthroughs*. Forecasting technological futures always inherently carries uncertainty. The technologies discussed herein are mature, and the degree to which China lags can be measured in well-defined generations through which other nations have had to progress, but some unlikely, radically novel way of producing advanced chips could theoretically emerge.

- *AAWs will require advanced chips.* Similarly, I argue that for various technical reasons, AAWs capable of substituting for human pilots and defeating modern air defenses will require advanced chips. While I believe the argumentation is sound, unexpected fundamental breakthroughs in algorithmic efficiency could theoretically mean "stupider" chips suffice. In such worlds, of course, there still might be decisive advantages to having the smartest chips, as I discuss below.

- *No peaceful reunification*. Taiwan Semiconductor Manufacturing Company (TSMC) is a critical node in the global semiconductor supply chain. While I argue below that China could not solve its

[426] Michael C. Horowitz, Sarah E. Kreps, and Matthew Fuhrmann, "Separating Fact from Fiction in the Debate over Drone Proliferation," *International Security* 41.2 (2016), 7-42; Michael C. Horowitz, Joshua A. Schwartz, and Matthew Fuhrmann, "China Has Made Drone Warfare Global," *Foreign Affairs*, November 20, 2020, www.foreignaffairs.com/articles/china/2020-11-20/china-has-made-drone-warfare-global.

[427] Shmuel Shmuel, "The Coming Swarm Might Be Dead on Arrival," *War on the Rocks*, September 10, 2018, https://warontherocks.com/2018/09/the-coming-swarm-might-be-dead-on-arrival/.

[428] Ibid.; Paul Scharre, "Counter-Swarm: A Guide to Defeating Robotic Swarms," *War on the Rocks*, March 31, 2015, https://warontherocks.com/2015/03/counter-swarm-a-guide-to-defeating-robotic-swarms/; Arthur Holland Michel, "Counter-Drone Systems, 2nd Edition," *Center for the Study of the Drone*, December 2019, https://dronecenter.bard.edu/files/2019/12/CSD-CUAS-2nd-Edition-Web.pdf.

[429] In this paper, by "advanced AI chips" I mean leading-node GPUs, FGPAs, and ASICs. I elaborate further below.

supply issues by forcibly seizing TSMC facilities during a conventional conflict, the story would look different if Taiwan had, years prior, voluntarily rejoined China in peacetime.

The argument unfolds in three parts. First, I briefly review the existing technological diffusion literature, including its predictions about AI. Second, I argue that even if AI *research* fully diffuses, lack of access to AI *hardware*, specifically advanced AI chips, would still prevent China from better leveraging AAWs in conventional conflict against the United States. Third, I show that China will not have access to such chips – their production, presently concentrated in the United States and allied countries, will not diffuse to China, making supply cutoff possible even after considering possible Chinese countermeasures.

## Existing Literature: Diffusion and AI

Scholars disagree about what existing theories of technological diffusion should apply to AI. In one camp, many IR scholars have argued beginning with Kenneth Waltz that the emulation of power-generating technologies possessed by other states is a mainstay of international affairs, following directly from international anarchy's demand that states secure their own survival.[430] Generally, scholars have emphasized that this emulation tends to eroding the leading state's position over time, as rising powers can free-ride off expensive R&D efforts and focus on implementation; for Robert Gilpin, this dynamic famously accounted for the cyclic rise and fall of hegemonic powers.[431] Analysts have submitted that globalization has likely accelerated the rate of diffusion of power, reducing the degree to which any state can be ahead of others.[432]

For another camp, however, the diffusion of military technology has arguably slowed. Most prominently, Andrea and Mauro Gilli argue that increased complexity has made advanced weapons systems, such as the US F-22, extremely difficult to copy, even given China's extensive skill at cyber-espionage. On their view, the pure engineering challenges now involved in assembling fighter aircraft, submarines, and aircraft carriers far exceed those faced by Imperial Germany in copying British dreadnoughts, for example.[433] Writing about current-generation drones, Gilli and Gilli also emphasize the increased "infrastructural support" needed for complex technologies, citing the long list of technologies involved in carrier strike groups other than carriers themselves. Contextually, using UAVs effectively requires integrating them into an organization capable of piloting them, processing the data received from them, and coordinating them with many other military assets.[434] The difficulty of these "infrastructural" tasks, in their view, has spiked.

---

[430] Kenneth Waltz, *Theory of International Politics* (New York: McGraw-Hill, 1979), 74, 127; Joao Resende-Santos, "Anarchy and the Emulation of Military Systems: Military Organizations and Technology in South America, 1870-1930," *Security Studies* 5 (1996): 193-260; Emily O. Goldman and Richard B. Andres, "Systemic effects of military innovation and diffusion*," Security Studies* 8 (1999).

[431] Alexander Gerschenkron, *Economic Backwardness in Historical Perspective: A Book of Essays* (MA: Belknap, 1962); Robert Gilpin, *War and Change in World Politics* (Cambridge: Cambridge University Press, 1981), 162, 176-85; Paul Kennedy, *The Rise and Fall of the Great Powers* (New York: Random House, 1987); George Modelski and William R. Thompson, *Leading Sectors and World Powers* (Columbia: University of South Carolina Press, 1996).

[432] Richard A. Bitzinger, "The Globalization of the Arms Industry: The Next Proliferation Challenge," *International Security* 19.2 (1994), 170-198; Moises Naim, *The End of Power: From Boardrooms to Battlefields and Churches to States, Why Being in Charge Isn't What It Used to Be* (New York: Basic Books, 2013); Thomas Friedman, *The World is Flat: A Brief History of the Twenty-first Century* (New York: Farrar, Strauss, and Giroux, 2005).

[433] Andrea Gilli and Mauro Gilli, "Why China Has Not Caught Up Yet: Military-Technological Superiority and the Limits of Imitation, Reverse Engineering, and Cyber Espionage," *International Security* 43.3 (2018/19): 141-89.

[434] Andrea Gilli and Mauro Gilli, "The Diffusion of Drone Warfare? Industrial, Organizational, and Infrastructural Constraints," *Security Studies* 25.1 (2016): 50-84. For the same argument contextualized to the Chinese Civil War,

Which of these theories better describes AI? The overwhelmingly predominant narrative holds that AI technology should easily and uniquely diffuse. After all, AI seems the child *par excellence* of today's "Information Age": as a general-purpose technology that goes in anything and everything, the commercial incentives to develop, spread, and integrate it are immense. The field itself reflects these expectations of maximum openness: cutting-edge research, including breakthrough results by leading companies, is almost universally posted on the unclassified internet in full methodological detail, simultaneous with the announcement of the results themselves. Unlike with spaceflight, the Internet, or GPS, no major government is pioneering basic research progress in AI; rather, the private sector is firmly ahead.[435] Overall, in the standard telling, one gets the impression that both state militaries and non-state actors might, in the near future, buy drones off Amazon, download the latest AI software from arXiv.org, and then easily combine them to produce weaponized swarms capable of terrorizing cities and seasoned human pilots alike. Given these conditions, combined with China's generally large efforts to acquire and adopt foreign technology, it seems hard to imagine that AI-enabled weapons would not diffuse to China.[436]

Indeed, in responding to the second camp, analysts tend to argue that AI may uniquely dodge these barriers: that is, AI's pattern of diffusion will more resemble smartphones, ubiquitous and commercialized, rather than F-35s, complex and rare. A 2020 Congressional Research Service report best encapsulates this view:

> AI systems are particularly vulnerable to theft by virtue of being almost entirely software-based ... the Chinese may be able to steal the plans for an F-35, but it will take them years to find the materials and develop the manufacturing processes to build one. In contrast, stolen software code can be used immediately and reproduced at will.[437]

Similarly, Michael Horowitz has argued that military AI systems may spread quickly if they have low unit costs and have core commercial in addition to military purposes, unlike the weapons technologies cited by Gilli and Gilli.[438] Horowitz, notably, also allows for the opposite possibility – if AI capabilities require

---

see Victor Cheng, "Modern War on an Ancient Battlefield: The Diffusion of American Military Technology and Ideas in the Chinese Civil War, 1946-1949," *Modern China* 35.1 (2009), 38-64.

[435] Lorand Laskai, "Civil-Military Fusion: The Missing Link Between China's Technological and Military Rise," *Council on Foreign Relations*, January 29, 2018, https://www.cfr.org/blog/civil-military-fusion-missing-link-between-chinas-technological-and-military-rise.

[436] Allison, "Is China Beating America to AI Supremacy?"; Brose, *The Kill Chain*, 90-117; Joseph S. Nye, Jr., *The Future of Power* (New York: PublicAffairs, 2011), 115; Miles Brundage, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, Paul Scharre, Thomas Zeitzoff, Bobby Filar, Hyrum Anderson, Heather Roff, Gregory C. Allen, Jacob Steinhardt, Carrick Flynn, Sean O hEigeartaigh, Simon Beard, Hadyn Belfield, Sebastian Farquhar, Clare Lyle, Rebecca Crootof, Owain Evans, Michael Page, Joanna Bryson, Roman Yampolsiky, and Dario Amodei, "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," *arXiv*, February 2018, https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf, 16-7; Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt, 2018), 81-104; Michael C. Horowitz, Gregory C. Allen, Elsa B. Kania, and Paul Scharre, "Strategic Competition in an Era of Artificial Intelligence," *CNAS*, 2018, https://www.cnas.org/publications/reports/strategic-competition-in-an-era-of-artificial-intelligence; William C. Hannas and Huey-meei Chang, "China's Access to Foreign AI Technology: An Assessment," *CSET*, 2019, https://cset.georgetown.edu/research/chinas-access-to-foreign-ai-technology/.

[437] Kelley M. Sayler, "Artificial Intelligence and National Security," *Congressional Research Service*, 2020, https://fas.org/sgp/crs/natsec/R45178.pdf, 34.

[438] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1.3 (2018), 45-6. This prediction derives from applying the theoretical framework, "adoption-capacity theory," from his seminal book, *The Diffusion of Military Power: Causes and Consequences for International Politics* (Princeton: Princeton University Press, 2010); see also William C. Wohlforth, Dmitry Adamsky,

classified data, pose difficult systems integration challenges, or threaten existing organizational hierarchies, they may diffuse less easily and grant innovators significant first-mover advantages.[439] For Allison, however, these may exactly be reasons to predict Chinese dominance – AI may threaten American organizational hierarchies, as well as antitrust and privacy-related values in American society more generally, but not the Chinese equivalents. In his words, "Beijing is not just trying to master artificial intelligence – it is succeeding."[440] America's competitors, the argument goes, will not bind their own hands with concerns about "AI ethics," whether domestically or on the battlefield.[441] In short, analysts mostly not only believe AI will easily diffuse, but that China may better utilize its fruits.

## Why AI Weapons Will Require Advanced Chips

In my view, however, analysts predicting diffusion have critically overlooked a key dimension: hardware. Although cutting-edge AI research is freely posted online, production of the specialized chips on which AI algorithms are best run is overwhelmingly concentrated in the United States and its allies. Below, I argue that even if Chinese AI research, organizational capital, and doctrine were all to overtake US capabilities, the United States would still be decisively advantaged in fielding AAWs if China were to lack access to advanced chips. The next section then argues that China will indeed lack access to such chips.

Notably, this is not an argument about the engineering complexity of future AI weapons, but rather an argument about the enforceable scarcity of critical material inputs. Analogously, possessing superior nuclear strategy and delivery systems would not be useful without fissile material. Given this, the best analogy for AI is not that "data is the new oil," but that "AI chips are the new uranium."[442] While uranium occurs ubiquitously in nature, enriching it to weapons-grade quality constrained proliferators in the early nuclear age. Similarly, while the raw materials for AI chips are everywhere, production of leading-edge chips is a highly rarefied ability. Even if AAWs are like iPhones, with both low unit costs and overwhelming overlap with various underlying commercial technologies, only Apple makes iPhones, and were they to cease selling them, consumers would face steep barriers in replicating the technology.[443]

I proceed in two parts: I first explain what "AI chips" are, and then explain why AAWs would almost certainly require them, including why substituting with larger numbers of older chips would fail.

---

Theo Farrell, Adam Grissom, Thomas G. Mahnken, and Michael C. Horowitz, "H-Diplo/ISSF Roundtable Review of Michael C. Horowitz. The Diffusion of Military Power: Causes and Consequences for International Relations (2010)," *Roundtable* 3.10 (2012), https://issforum.org/ISSF/PDF/ISSF-Roundtable-3-10.pdf.

[439] Ibid., 52-3.

[440] Allison, "Is China Beating America to AI Supremacy?"

[441] A large literature argues this point. See Brose, *The Kill Chain*, 114-7; Lee, *AI Superpowers*; 81-103; Tim Hwang, "Shaping the Terrain of AI Competition," CSET, 2020, https://cset.georgetown.edu/research/shaping-the-terrain-of-ai-competition/; H. Akın Unver, "Artificial Intelligence, Authoritarianism and the Future of Political Systems," *Center for Economics and Foreign Policy Studies*, 2018; Jinghan Zeng, ""Artificial intelligence and China's authoritarian governance," *International Affairs* 96.6 (2020): 1441-59.

[442] "Data is the new oil" has other conceptual problems. Most glaringly, data is not fungible between tasks – surveillance data on Chinese citizens is not useful for training air-to-air combat algorithms for AAWs, for example. For further discussion, see the first paper of this dissertation.

[443] Gilli and Gilli, "Correspondence: Military-Technological Imitation and Rising Powers," 191.

## (1) What Are AI Chips?

Unlike more universally used CPUs, AI chips have specific design features optimized for implementing AI models, such as greater ability to execute highly parallel calculations, deliberately lower-precision calculation of numbers sufficient for AI models but more efficient in use of transistors, and/or adaptation to programming languages specifically written for maximizing AI code efficiency. Consequently, specialized AI chips can provide improvements over CPUs equivalent to 26 years of CPU development. Further, even AI chips which are specialized but not state-of-the-art can incur order-of-magnitude greater energy costs.[444]

Moore's Law, which states that the number of transistors per chip doubles every two years, held from the 1960s to the 2010s, then began slowing as chip design ran up against fundamental physical limits.[445] Individual transistors are now only several atoms thick, posing increasingly esoteric challenges; in the 2000s, for example, engineering difficulties arose when electrical current began leaking out from between atomically thin insulative layers of transistors.[446] Consequently, further progress has involved specialized chips, rather than simply cramming more transistors into CPUs. There are three types of AI chips:

- GPUs ("graphics processing units"), originally designed for image-processing applications, exhibit the ability to execute highly parallel computation. GPUs became the most-used chip for training AI systems in 2017.[447]

- FPGAs ("field-programmable gate arrays"), as the name suggests, include logic blocks which can be programmed after the FPGA's manufacture to optimally execute specific algorithms.[448]

- ASICs ("application-specific integrated circuits"), in contrast with FPGAs, are *hardwired* to optimally execute specific algorithms. State-of-the-art ASICs beat their FPGA equivalents, but become obsolete quickly since each new algorithm requires new ASICs to be manufactured, while FPGAs can be reprogrammed.[449]

## (2) Are AI Chips Necessary?

Accepting that "AI chips" are better for AI, why would advanced AI chips specifically be necessary for AAWs? In short, increasingly complex AI models will pose escalating difficulties with both computation and miniaturization, gating their ability to be updated and installed on weapons platforms intended to navigate battlefield environments autonomously. The process of using AI systems can be divided into two distinct phases: training and inference. I discuss each phase in turn.

---

[444] Saif M. Khan and Alexander Mann, "AI Chips: What They Are and Why They Matter," *CSET*, April 2020, https://cset.georgetown.edu/wp-content/uploads/AI-Chips%E2%80%94What-They-Are-and-Why-They-Matter.pdf, 23, 5-6.

[445] Ibid., 7.

[446] Ibid., 10.

[447] Ibid., 20.

[448] Ibid.

[449] Khan and Mann, "AI Chips," 20.

### (a) Training

First, one undertakes training of the AI model, during which it learns the knowledge to be used later in the field. For example, training AAWs might involve showing them many pictures of munitions fired by enemy air defenses, perhaps combined with simulated practice in a virtual environment, so as to teach them to evade effectively. Axiomatically, viable use of drones in interstate warfare will require more advanced machine learning models than presently fielded – existing autonomous drones cannot defeat modern air defenses; improvements in navigation, evasiveness, target selection, and other metrics will be required to begin replacing more of the functions of human pilots.[450]

Importantly, these more advanced models will likely require exponentially more computing power.[451] The amount of computational power used in training the most sophisticated AI models has, since 2012, doubled every 3-4 months. Consequently, compute usage has increased 300,000 times since 2012, driven significantly by deep learning's extreme responsiveness to the use of ever greater quantities of compute.[452] As Rich Sutton, the father of modern reinforcement learning, pithily remarked, the last 70 years of AI research have shown that leveraging greater computing power beats clever engineering.[453] Similarly, analyzing 171,394 deep learning papers, Nur Ahmed and Muntasir Wahed find that the necessity of computing power for deep learning has meant large firms and elite universities dominate the AI frontier.[454] Compute, not data, bottlenecks the modern development of AI models.

Thus, whenever battlefield-effective AAW algorithms first become available, they will likely first result from quantities of compute expenditure only available to the largest AI companies and/or most AI-invested militaries. Further, as battlefield conditions change, maintaining updated AAW algorithms will demand continuous training of AI models. Using deep learning, AI training processes often involve executing the same computation millions of times; intuitively, this means the parallelism-optimized architectures of advanced AI chips benefit the speed and efficiency of such processes enormously.[455] Thus, lack of access to leading-edge server-grade chips would necessarily constrain the quality of Chinese AAW algorithms used for swarming, navigation, evasiveness, and any other complex cognitive functions currently performed by human pilots. According to the chip benchmarking literature, which studies the relative efficacy of different chip types for AI applications, specialized AI chips exhibit up to three orders of magnitude of improvement, relative to even state-of-the-art CPUs, in efficiency and speed across both training and inference tasks.[456] If it were to take 1,000 times as long for China to produce updated AAW algorithms,

---

[450] Zegart, "Cheap fights, credible threats"; Gilli and Gilli, "The Diffusion of Drone Warfare?", 50-84; Michael C. Horowitz, Sarah E. Kreps, and Matthew Fuhrman, "The Consequences of Drone Proliferation: Separating Fact from Fiction," *International Security* 41.2 (2016): 7-42.

[451] Dario Amodei and Danny Hernandez, "AI and Compute," *OpenAI*, May 16, 2018, https://openai.com/blog/ai-and-compute/.

[452] Ibid.

[453] Rich Sutton, "The Bitter Lesson," *Incomplete Ideas*, March 13, 2019, http://www.incompleteideas.net/IncIdeas/BitterLesson.html.

[454] Nur Ahmed and Muntasir Wahed, "The De-Democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research," *arXiv*, October 22, 2020, https://arxiv.org/ftp/arxiv/papers/2010/2010.15581.pdf.

[455] Khan and Mann, "AI Chips," 19.

[456] Ibid., 23, 38-40.

this would obviously dampen their efficacy. Older or unspecialized chips may simply be unable to be used for the same applications, as energy consumption would "quickly balloon to unmanageable levels."[457]

Notably, even assuming perfect espionage, China would not be able to field equivalently effective AAWs simply by stealing already-trained US AI models. After all, US and Chinese AAWs would be optimized for different tasks, priorities, and modes of operation, given each country's differing objectives, geography, and orders of battle. For example, analysts often derive Chinese advantage from the idea of using AAWs to overwhelm US aircraft carriers. The US itself would have a greatly reduced need to train AI models which, when loaded onto AAWs, generated optimum anti-carrier tactics, however. Thus, models most suited for China's specific strategic purposes may simply not be available to steal. Even when reaching the state-of-the-art is still possible, training runs using lower-quality chips can cost orders of magnitude more money, due to relative inefficiency. Thus, this would also dampen a major theorized tactical advantage of AAWs relative to high-tech manned platforms, which is their relative cheapness.[458]

**(b) Inference**

Further, apart from difficulties with training, advanced AI chips will also be necessary for inference – actual application of the trained AI model in the field. Specifically, it likely will not be possible to deploy appropriately miniaturized AAWs without state-of-the-art chips. The reason is the breakdown of Dennard scaling, a counterpart of Moore's Law, which held that power density would remain constant even as transistor count increased for some unit area.[459] Since this is no longer true as of the 2000s, advanced AI models will be constrained in their ability to naturally become able to be loaded onto smaller and smaller packages over time, as the heat generated by computation would destroy the platform. Put another way, weapons platforms would have to be two or three orders-of-magnitude physically larger to run the same deep learning applications, if state-of-the-art chips are not used; in addition, due to increased power consumption overall, they would possess dramatically less range.[460] Illustratively, to return to the above-invoked analogy, the degree to which uranium has been enriched determines the quantity required to build a nuclear weapon, as less is required at more enriched levels to reach critical mass. At 5% enrichment, critical mass would require an elephant-sized quantity; weapons-grade uranium is typically defined as 90% or higher, requiring only about a turkey's worth to go critical.[461] Very similarly, the higher quality the chips used, the smaller the drones can be. Thus, future drones able to load sufficiently complex AI models to autonomously swarm or defeat modern air defenses will almost certainly require advanced AI chips.[462]

---

[457] Ibid., 23-6.

[458] Saif M. Khan, "Maintaining the AI Chip Competitive Advantage of the United States and its Allies," *CSET*, 2019, https://cset.georgetown.edu/wp-content/uploads/CSET-Maintaining-the-AI-Chip-Competitive-Advantage-of-the-United-States-and-its-Allies-20191206.pdf, 3.

[459] The seminal paper is Hadi Esmaeilzadeh, Emily Blem, Renée St. Amant, Karthikeyan Sankaralingam, and Doug Burger, "Dark silicon and the end of multicore scaling" (paper presented at the 38th Annual International Symposium on Computer Architecture, San Jose, CA, 2011), 365-376. A readable explanation is available at Silvano Gai, "Dennard Scaling," March 17, 2020, https://silvanogai.github.io/posts/dennard/. For a short, technical explanation, see Mark Bohr, "A 30 Year Retrospective on Dennard's MOSFET Scaling Paper," 2007, *Solid-State Circuits Society*, http://www.eng.auburn.edu/~agrawvd/COURSE/READING/LOWP/Boh07.pdf.

[460] Tim Hwang, "Computational Power and the Social Impact of Artificial Intelligence," *MIT Media Lab*, 2018, 15-6. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3147971.

[461] "Uranium Enrichment," *Nuclear Threat Initiative*, accessed 2021, https://tutorials.nti.org/nuclear-101/uranium-enrichment/.

[462] Hwang, "Computational Power and the Social Impact of Artificial Intelligence," 15-6. Another useful analogy may be jet fuel: the first aircraft used the same gasoline burned by automobiles. When military jet engines were developed

### (c) Could China use a cloud computing architecture?

One could ask: why would it be necessary to perform inference on AAWs themselves (an "edge computing" architecture), as opposed to, say, beaming the data back to an on-shore data-center (a "cloud computing" architecture)? A cloud architecture would face several issues specific to the intended functionality of AAWs – first, most obviously, this would introduce a significant vulnerability, as the signal could be jammed or spoofed.[463]

Second, this would introduce severe issues related to latency. How long could AAWs wait, after all, if attempting to perceive and dodge anti-aircraft fire? Some evidence about this dynamic is available from development of autonomous cars, which rely on AI chips for high inference speed, since they must rely quickly to changing traffic conditions (e.g., to avoid hitting pedestrians).[464] For AAW-on-AAW conflicts, edge computing would almost certainly be necessary to be competitive – that is, since mere seconds of maneuver decisions can decide the outcomes of dogfights, AAWs with better AI chips performing faster inference will naturally be able to defeat adversaries by more rapidly executing the so-called OODA ("Observe, Orient, Decide, Act") loop often used to characterize aerial combat.[465] Even if China located AAW servers on its eastern coast to minimize distance to likely battlefields, even a few seconds would likely systematically disadvantage its AAWs compared to those performing on-board computation.[466]

Relatedly, specialized AI chips can further increase computational speed through architectural innovation: for example, US-based Google's TPU ("tensor-processing unit"), an AI-specialized ASIC, contains sufficient on-chip memory to run AI models on the chip itself. In contrast, forcing the chip to communicate with external memory in other parts of the platform can take, depending on the details, up to 100 times longer.[467] Thus, hardware progress would likely continue to advantage AAWs with more advanced AI chips over those with less.

### (d) Could China substitute with larger numbers of older chips?

The reader may intuitively object: could China simply use larger numbers of worse chips? If China can simply substitute larger numbers of lower-quality chips for AAWs, after all, then lack of access to cutting-edge chips may become only a marginal economic cost, rather than a decisive supply line interaction. Even for training, several additional technical reasons count against this possibility, however.

---

in the leadup to World War II, however, more refined fuels with higher flashpoints were required to avoid unacceptable rates of accidental fire. Consequently, the US military developed specialized mixtures which became the first jet fuels. Similarly, current-generation drones can use run-of-the-mill computer chips, but AAWs may require the "jet fuel" of advanced AI chips. See "Aviation Fuel," *U.S. Centennial of Flight Commission*, accessed April 25, 2021, https://web.archive.org/web/20120420064213/http://www.centennialofflight.gov/essay/Evolution_of_Technology/fuel/Tech21.htm.

[463] For similar concerns about hacking, US nuclear weapons do not allow in-flight retargeting.

[464] Khan and Mann, "AI Chips," 21.

[465] Michael C. Horowitz, "When speed kills: Lethal autonomous weapon systems, deterrence and stability," *Journal of Strategic Studies* 42 (2019): 764-88.

[466] And, of course, these servers would then also represent fixed targets vulnerable to attack.

[467] See Gaurav Batra, Zach Jacobson, Siddarth Madhav, Andrea Queirolo, and Nick Santhanam, "Artificial-intelligence hardware: New opportunities for semiconductor companies," *McKinsey*, 2019.

First, as mentioned, leading AI algorithms depend on handling massive data through model parallelism – that is, they break down the tasks required for running AI into many smaller, simultaneous processes. Difficulties in tuning for parallelism, however, scale with the number of chips intended to be used in parallel. Consequently, whenever AAWs able to compete on the modern battlefield emerge as a breakthrough capability, the ability to implement their algorithms on larger numbers of older chips will be a distinct, unsolved technical problem. This problem is not likely to be small, nor is it one that is commonly solved, as leading AI companies tend to simply acquire leading-node chips. Depending on the algorithmic details, implementation on large numbers of older chips may even simply be technically impossible.[468]

Thus, if China lacked access to advanced chips, it could not simply copy-paste American AAW algorithms, even if stolen *in toto* through espionage. Rather, it would then have to confront an additional, difficult, and potentially impossible software engineering problem of adapting the stolen algorithms to older hardware. Since inefficiency in power usage means the operating costs of using old chips exceeds leading-node AI chips substantially, this could theoretically even make attrition-style AAW air-to-air combat no longer a favorable cost-exchange proposition for China.[469]

Second, even assuming China could solve this engineering problem, it would then also require additional "networking technology," which governs how quickly and efficiently chips can "talk" to other chips, to implement. Otherwise, the increased volume of inter-chip communication required by using more chips would necessarily mean slower inference speeds, allowing American AAWs to outmaneuver Chinese copies in executing the OODA loop. Thus, China would face both an additional hardware problem, in addition to the above algorithmic problem.[470]

Finally, it is difficult to overstate how impoverished China's indigenous chip fabrication capability is – China does not have the ability to fabricate "chips that are nearly as good as the state-the-art," with which it could conceivably substitute for leading-edge chips; rather, even considering AI chips more broadly, up to 16-nanometer transistors and better, the US, Taiwan, and South Korea still control 95.3% of global fabrication capability.[471] I discuss China's shortcomings in indigenous production in the next section.

## No Chips for the Drones of China

Even if advanced AI chips will be required for AAWs, can access to them be denied to China? I first show why production of advanced AI chips will not diffuse away from the United States and its allies, then explain why a supply cutoff would be effective.

### (1) Why Hardware Production Will Not Diffuse

The United States and its allies currently control, and are likely to control, the production of leading-node AI chips for at least the next decade. This section proceeds in two parts: First, I explain what theoretical mechanisms have driven why the AI supply chain has become increasingly concentrated, contrary to scholarly predictions about globalization. Second, I show that this concentration is currently controlled by the United States and its allies, and argue that this state of affairs will persist for at least the next decade.

[468] Khan and Mann, "AI Chips," 27.

[469] Ibid., 25.

[470] Ibid., 27.

[471] Khan and Flynn, "Maintaining China's Dependence on Democracies for Advanced Coputer Chips," 5.

### (a) In AI, globalization means specialization, not redundancy

Naively, since AI chips have intense commercial value, one might expect their production to diffuse widely. For example, Horowitz assesses that AI may "decentralize the capacity to produce important military hardware," as weapons systems increasingly based on dual-use commercial technologies should become widely available. Thus, although China may find it very difficult to overtake the United States "in the area of manned fighters, manned tanks, or aircraft carriers," conditions may change if "the shift to systems based more on commercial technologies … allow[s] more countries to produce military power independently than could before."[472] With AI, however, commercialization has actually resulted in the opposite trend – *fewer* countries can produce power independently, because particular segments of the involved technologies have become increasingly concentrated in fewer and fewer companies and countries.

Fundamentally, this is a result of specialization: ironically, as I show in the next section, the reliability of global trade in the AI supply chain has encouraged not only specialization according to comparative advantage, but also the pushing of scientific limits through pooling global demand and talent which would not have been possible in autarkic countries. That is, according to a broad swath of evidence across sectors of the global economy, research productivity appears to have gradually and broadly declined over the last century, whether considering agricultural productivity, medical innovations, the price of light, genomic science, or many other fields.[473] With AI specifically, the number of researchers used as inputs to double chip density has increased by 18 times since the early 1970s.[474]

Research fields have generally reacted to increased R&D difficulty by combining global forces, and advanced AI chip production has been no exception. Surviving firms in AI require both almost all available global talent, as well as the economy of scale created by meeting almost all global demand, to make continuous R&D progress. In 2000, 25 chip manufacturers made cutting-edge chips; in 2018, only 3 did.[475] The fixed costs required to stay competitive in semiconductor manufacturing have repeatedly ballooned, encouraging iterated consolidation or relinquishment. For example, the average AI chip company spends 25% of revenue on R&D.[476] Modern fabrication plants ("fabs") cost billions of dollars, and cutting-edge chips each contain tens of billions of transistors.[477] From 2017-2020, Samsung spent $93.2 billion in semiconductor capital outlays; in 2021, Taiwan's TSMC announced a $100 billion investment over three years, a figure exceeding the GDP of two-thirds of the world's countries, to expand capacity.[478] Designing

---

[472] Michael C. Horowitz, "Information-Age Economics and the Future of the East Asian Security Environment," in *The Nexus of Economics, Security, and International Relations in East Asia*, eds. Avery Goldstein and Edward D. Mansfield (US: Stanford University Press, 2012), 211-2, 223-4, 227.

.

[473] This is a large literature. For the best overview, see Nicholas Bloom, Charles I. Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?", *American Economic Review* 110.4 (2020: 1104-144. For an earlier overview, see Benjamin F. Jones, "The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder?" *Review of Economic Studies* 76.1 (2009), 283–317.

[474] Bloom et al., "Are Ideas Getting Harder to Find?," 1105.

[475] I visualize this consolidation in logic chip manufacture further below.

[476] Kleinhans and Baisakova, "The global semiconductor value chain," 7.

[477] Ibid., 14-5.

[478] "Samsung and TSMC Seeking to Spend Their Way to Worldwide Domination of Advanced IC Technology," IC Insights, March 16, 2021, https://www.icinsights.com/news/bulletins/Samsung-And-TSMC-Seeking-To-Spend-Their-Way-To-Worldwide-Domination-Of-Advanced-IC-Technology/; Robert Clark, "Samsung, TSMC $130B plunge underlines the great chip divide," Light Reading, 2021, https://www.lightreading.com/asia/samsung-tsmc-$130b-plunge-underlines-great-chip-divide/d/d-id/768515.

a 10-nanometer node chip cost $170 million in 2016, while designing a 5nm node chip cost $540 million in 2020.[479] EUV machines, which cost $160 million per unit, print transistors with precision equivalent to about the growth of a human fingernail in a few seconds.[480] In all, producing a single chip involves more than 1,000 individually complex steps.[481] Further, since deep learning benefits greatly from specialized AI chips, its continuously increasing prominence in modern AI is only likely to further accelerate this trend.[482]

In sum, commercialization and globalization have disincentivized indigenous production, and instead driven a tendency toward an increasingly specialized, fragmented number of distinct market segments, which either begin or become more concentrated as they arise and develop over time. As a practical consequence of these trends, across the AI supply chain, due to the economies of scale, concentrations of engineering talent, and volumes of tacit knowledge required to sustain these efforts, for many highly specific sub-technologies only a small handful of companies or sometimes only one company can exist, with even highly aggressive state subsidies empirically unable to sustain cutting-edge capability domestically. Leveraging industry datasets, I illustrate this for various parts of the AI supply chain below.

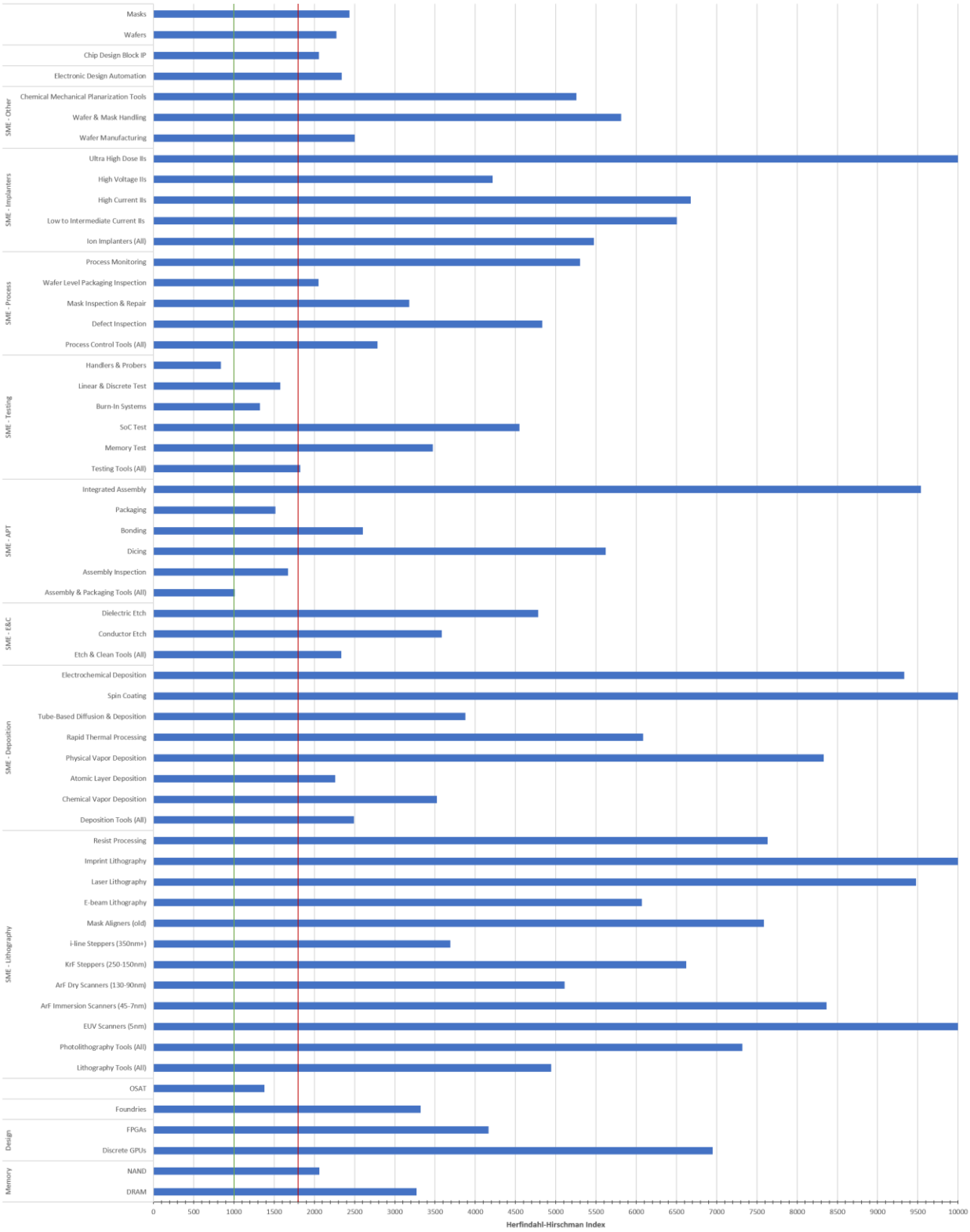
[479] Kleinhans and Baisakova, "The global semiconductor value chain," 12.

[480] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 8.

[481] Khan et al., "The Semiconductor Supply Chain," 5.

[482] Neil C. Thompson and Svenja Spanuth, "The Decline of Computers as a General Purpose Technology: Why Deep Learning and the End of Moore's Law are Fragmenting Computing," SSRN, 2018 (revised version forthcoming in Communications of the ACM), http://ide.mit.edu/sites/default/files/publications/SSRN-id3287769.pdf.

## Herfindahl-Hirschman Index – All Segments[483]




[483] Data sources: SEMI; VLSI Research; TrendForce; public reporting.

Specifically, the horizontal axis is the segment's Herfindahl-Hirschman Index (HHI) score, a standard measure of market concentration, illustrating declining firm count at the cutting-edge of various sub-technologies.[484] HHI is calculated by adding the square of the market share of each firm within a given market. Consequently, under HHI, 10,000 is a perfect monopoly. A market under 1,000 is coded by the US Fair Trade Commission as unconcentrated; between 1,000 and 1,800, the market is "moderately concentrated"; above 1,800, the market is "highly concentrated."[485] Below, I draw a green line at 1000, and a red line at 1800 – as is visually apparent, the AI supply chain is overall extremely concentrated.

These trends toward concentration can also be understood by examining specific segments over time. For example, as an oft-cited example of gradual monopoly, I diagram below the progressive dropping-out of companies in the photolithography segment, eventually leaving Dutch company ASML with a monopoly over EUV photolithography technology, required to produce all leading-edge AI chips.

**Photolithography – Consolidation Over Time**[486]

| | | Netherlands | Japan | | | US | |
|---|---|---|---|---|---|---|---|
| | | ASML | Nikon | Canon | SVGL | Ultratech | Perkin Elmer |
| i-line | 1990 | Achieved | Achieved | Achieved | Achieved | Achieved | Achieved |
| | 1994 | Achieved | Achieved | Achieved | Achieved | Achieved | Dropped Out |
| Krypton flouride (KrF) | 1995 | Achieved | Achieved | Achieved | Achieved | Dropped Out | |
| | 1997 | Achieved | Achieved | Achieved | Achieved | | |
| | 1999 | Achieved | Achieved | Achieved | Achieved | | |
| | 2001 | Achieved | Achieved | Achieved | Dropped Out | | |
| Argon flouride (ArF) | 2004 | Achieved | Achieved | Dropped Out | | | |
| | 2006 | Achieved | Achieved | | | | |
| Argon flouride immersion (ArFi) | 2009 | Achieved | Achieved | | | | |
| | 2011 | Achieved | Achieved | | | | |
| | 2014 | Achieved | Achieved | | | | |
| | 2015 | Achieved | Achieved | | | | |
| | 2017 | Achieved | Achieved | | | | |
| | 2018 | Achieved | Achieved | | | | |
| Extreme ultraviolet (EUV) | 2020 | Achieved | Dropped Out | | | | |

Similarly, we could also consider specific AI chip types – again, GPUs, FPGAs, and ASICs. Nvidia invented discrete GPUs in 1999, for example, and only Nvidia and AMD have had meaningful market share through at least 2019.[487] Unfortunately, ASIC company revenues are not publicly available, but the below figure shows two snapshots of the FPGA segment for which data are available: 1996 and 2016. By 2016, the market became dominated by three major companies.

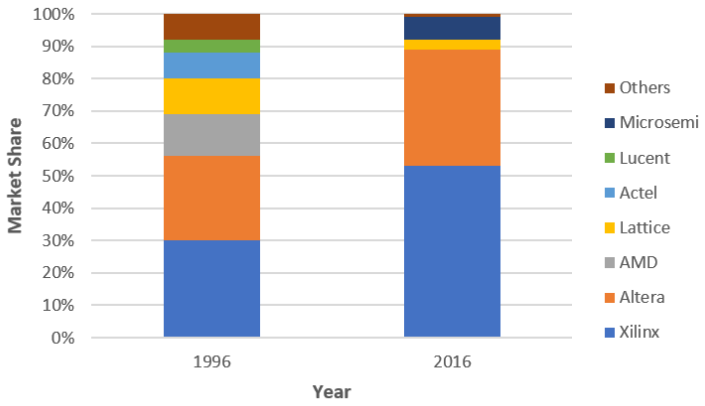
[484] In studying oil, Hughes and Long (2014/15) similarly use this metric as a proxy variable for coercive power across market segments. See Llewelyn Hughes and Austin Long, "Is There an Oil Weapon? Security Implications of Changes in the Structure of the International Oil Market," *International Security* 39.3 (2014/15): 152-89.

[485] Ibid.

[486] Adapted from Will Hunt, Saif Khan, and Dahlia Peterson, "China's Progress in Semiconductor Manufacturing Equipment: Accelerants and Policy Implications," *CSET*, March 2021, 10-11.

[487] "AMD and NVIDIA Add-in-Board GPU Market Share from 2002 to Q3/2016," *TechPowerup*, 2016, https://www.techpowerup.com/228095/amd-and-nvidia-add-in-board-gpu-market-share-from-2002-to-q3-2016; Hassan Mujtaba, "AMD Radeon GPUs Gained Big Chunk Of Discrete Market Share Versus NVIDIA GeForce In Q4 2019 – Up From 27% To 31%," *WCCFTech*, 2020, https://wccftech.com/amd-radeon-and-nvidia-geforce-discrete-gpu-market-share-q4-2019/.

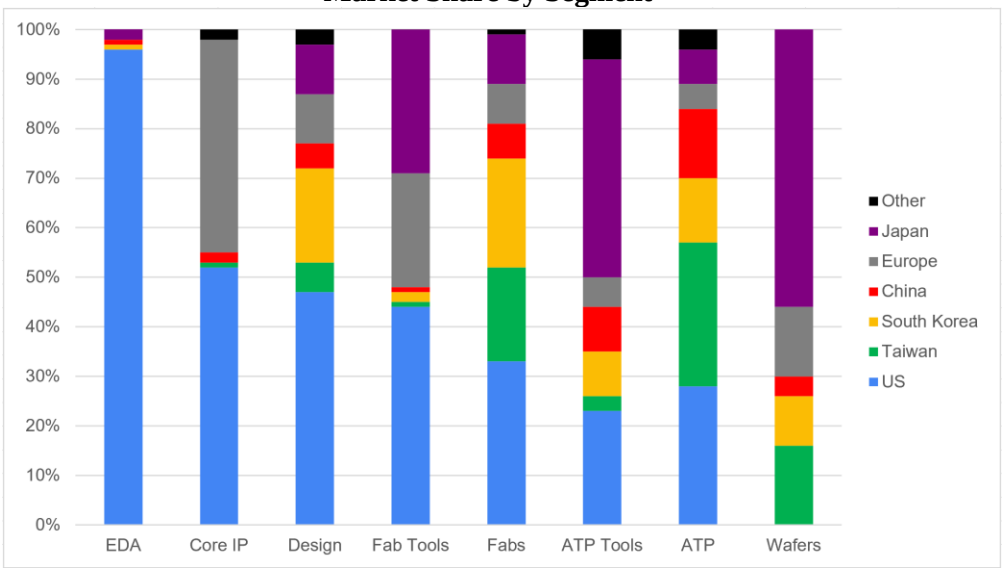
**FPGAs – Consolidation Over Time**[488]




**(b) AI segments live on US and allied soil**

Further, this consolidation has overwhelmingly led to concentrating AI supply chain segments in the United States and allied countries. While Horowitz writes that both China and Taiwan "lead the world" in semiconductors, the two countries are in fact worlds apart – China lags in almost every segment.[489] Even at the broad cross-segment level, the US and allied advantage is easily visible.

**Market Share by Segment**[490]



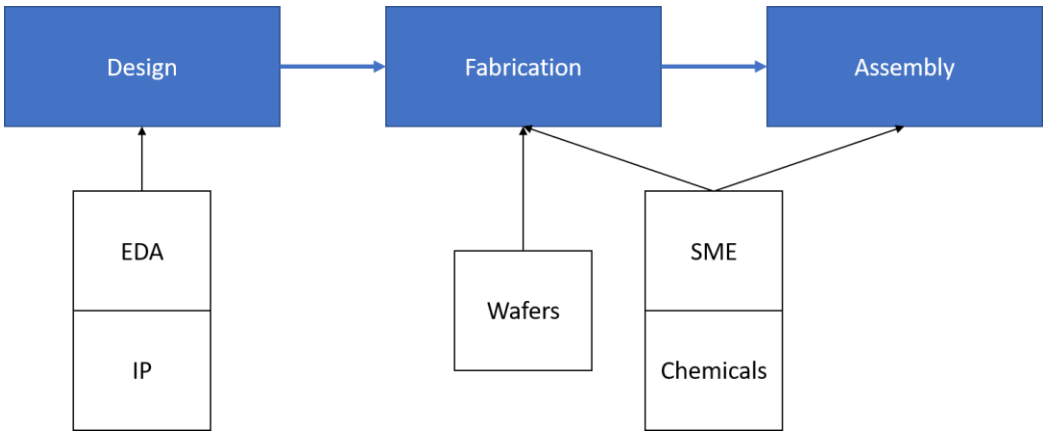

[488] Data sources: Hwang, "Computational Power and the Social Impact of Artificial Intelligence"; "Analysis of Xilinx's Patent Strategies," *Berkeley EECS*, accessed 2021, https://inst.eecs.berkeley.edu/~eecsba1/sp98/reports/eecsba1a/p2frame.html.

[489] Horowitz, "Information-Age Economics and the Future of the East Asian Security Environment," 213.

[490] Data source: Saif M. Khan, Alexander Mann, and Dahlia Peterson, "The Semiconductor Supply Chain: Assessing National Competitiveness," *CSET*, 2021, https://cset.georgetown.edu/wp-content/uploads/The-Semiconductor-Supply-Chain-Issue-Brief.pdf, p. 8.

Further, it may also be instructive to walk, step-by-step, through the supply chain in somewhat more detail, and examine concentration at select parts thereafter.

**The Semiconductor Supply Chain**[491]




At a high level, the production process for chips has three steps: design, fabrication (i.e., manufacture), and assembly. Chip design makes use of chip IP and electronic design automation ("EDA") software, which form separate market segments; fabrication creates chips out of wafers (most frequently silicon), and both the fabrication and assembly steps use various chemicals and semiconductor manufacturing equipment ("SME"). While most companies previously performed all three steps in-house, modern-day companies now often specialize, owing to increasing complexity.[492]

**Design**

In the design of discrete GPUs, US-based Nvidia and AMD have a duopoly.[493] While ASIC data is unfortunately unavailable, US-based Xilinx and Intel dominate FPGA design, while Chinese FPGA companies have only developed trailing-node chips.[494] Further, all chip design companies require access to EDA software. There are essentially only three EDA companies, all US-based: Cadence Design Systems, Synopsys, and Mentor. EDA was dominated by these three companies in 2001; this remained true in 2018.[495] EDA is highly specific to particular fabs and nodes, meaning that even successfully stealing previous-generation EDA would only grant limited and decaying usefulness. EDA vendors spend 35% of revenue on R&D, reflecting the complexity of their industry. [496] In contrast, China has less than 1% market

[491] Adapted from Kleinhans and Baisakova, "The global semiconductor value chain," 12.

[492] The notable exceptions are Intel and Samsung, although Intel may be shifting away from an integrated model due to encountering various technical difficulties. See Leo Sun, "Why Intel's Foundry Plans Don't Make Any Sense," *The Motley Fool*, 2021, https://www.fool.com/investing/2021/05/03/why-intel-foundry-plans-dont-make-any-sense/.

[493] Khan and Mann, "AI Chips," 27-8.

[494] Ibid.

[495] Cheng Ting-Fang and Lauly Li, "The great US-China tech decoupling: Where are we now?," *Nikkei Asia*, 2019, asia.nikkei.com/Economy/Trade-war/The-great-US-China-tech-decoupling-Where-are-we-now; Russ Arensman, "EDA's OK," *EDN*, 2002, https://www.edn.com/edas-ok/.

[496] Kleinhans and Baisakova, "The global semiconductor value chain," 13.

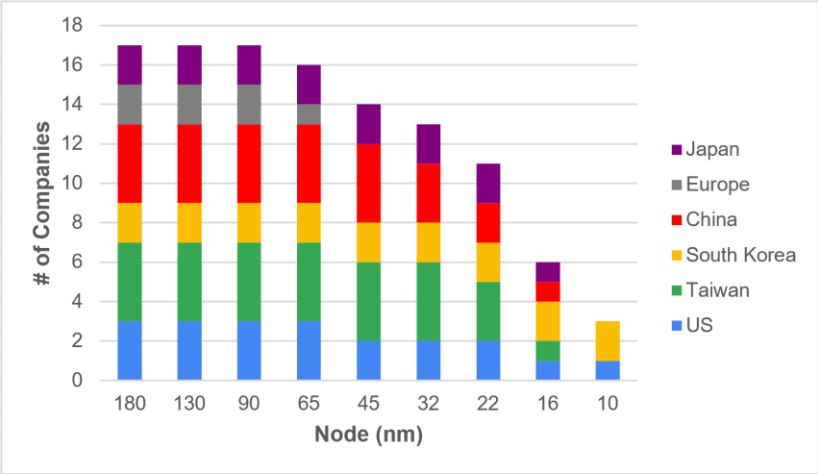
share of the global chip design market, including high-end CPUs, discrete GPUs, FPGAs, and memory chips.[497]

**Fabrication**

Let us first discuss memory chips, which function as storage, and then logic chips, which undertake computational processing. Beginning with memory, DRAM ("dynamic random-access memory") chips are used in all computing devices: specifically, they are "short-term memory" chips which temporarily store data while processing it. The DRAM market has consolidated substantially: in 2005, the eight largest DRAM vendors had 97% market share; in 2019, the three largest DRAM vendors had 95% market share. In other words, the DRAM market has consolidated into three major players: South Korea's Samsung and SK Hynix, and US-based Micro.[498]

On the other hand, NAND flash memory chips, named after the NOT-AND logic gate, are used in most computing devices: specifically, they are the "long-term memory" counterpart to DRAM, storing data for the longer-term. NAND is dominated by six major players: South Korea's Samsung and SK Hynix, Japan's KIOXIA, and US-based Micron and Intel.[499] However, only US and Korean companies have made it to the 10nm node.

**Memory Chipmakers by Country and Node**[500]




Moving to logic fabs, which produce all "AI chips," one notes that they are extremely concentrated, especially at the leading edge. Taiwan's TSMC has more than 50% market share, followed by South Korea's Samsung at a distant second. TSMC and Samsung are the only fabs offering fabrication at 7nm nodes or better - while US-based Intel also plans to do so, it stumbled in transitioning nodes and now plans to have 7nm production online only in 2023, assuming no further delays.[501]

---

[497] Khan et al., "The Semiconductor Supply Chain," 16-7.

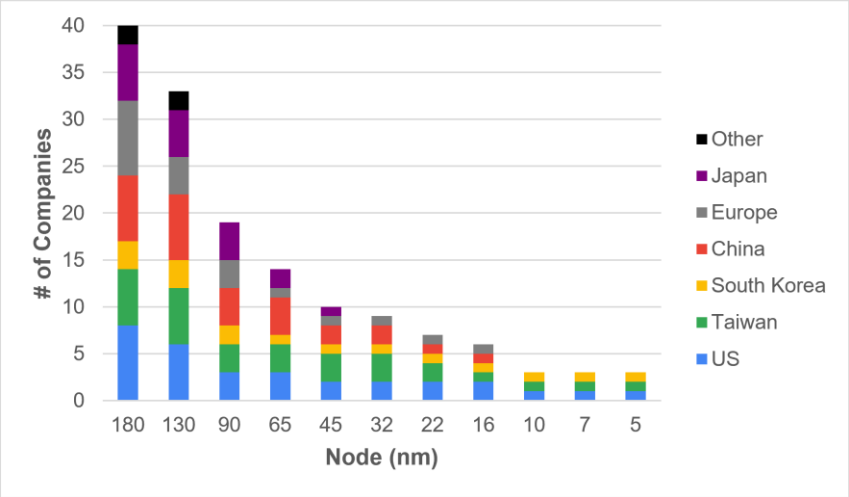
[498] Ibid., 7-8.

[499] Ibid.

[500] Source: SEMI, November 2020.

[501] Leo Sun, "Why Intel's Foundry Plans Don't Make Any Sense," *The Motley Fool*, 2021, https://www.fool.com/investing/2021/05/03/why-intel-foundry-plans-dont-make-any-sense/.

**Logic Chipmakers by Country and Node**[502]




While the Chinese government has heavily subsidized its own fab, SMIC, it heavily trails TSMC and over 90% of its business is older 250-40nm nodes.[503] State subsidies were up to a whopping 40% of total chip fab revenue for Chinese firms from 2014-2018, and yet the most heavily subsidized fab, SMIC, was only able to produce a limited amount of 14nm capacity.[504] One industry expert remarked, "I think what they are doing in fabrication is another Great Leap Forward."[505] Despite disproportionate government equity injections, Chinese firm profitability remains behind European, Japanese, Korean, Taiwanese, and American peers.[506]

**Semiconductor Manufacturing Equipment**

Next, as mentioned above, all fabs rely on semiconductor manufacturing equipment (SME) companies for the equipment they use to produce chips. SME itself consists of many specialized and consolidated segments, with vendors specializing in equipment used in specific steps out of hundreds of the chip fabrication process. Five companies dominate the SME market: US-based Applied Materials, Lam Research, and KLA; the Netherlands' ASML, and Japan's Tokyo Electron.[507] Incumbents are very difficult to displace because they receive continuous feedback and knowledge from their collaborations with, and

[502] Source: SEMI, November 2020.

[503] Kleinhans and Baisakova, "The global semiconductor value chain," 15.

[504] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 4.

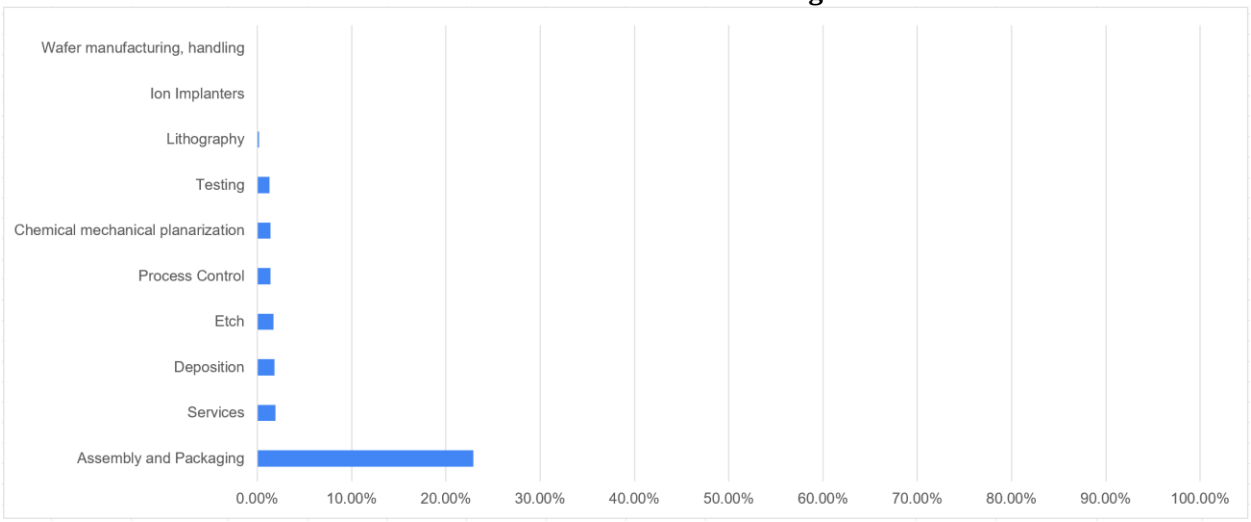
[505] Douglas B. Fuller, "Growth, Upgrading, and Limited Catch-Up in China's Semiconductor Industry," in *Policy, Regulation and innovation in China's Electricity and Telecom Industries*, eds. Loren Brandt and Thomas G. Rawski (Cambridge University Press, 2019), 297.

[506] Andrea Andrenelli, Julien Gourdon, Yuki Matsumoto, Taku Nemoto, Jehan Sauvage, and Christian Steidl, "Measuring distortions in international markets: The semiconductor value chain," *OECD Trade Policy Papers* 234 (2019), https://www.oecd-ilibrary.org/docserver/8fe4491d-en.pdf?, 73.

[507] Kleinhans and Baisakova, "The global semiconductor value chain," 16-7.

sales to, leading fabs.[508] Together, the US, the Netherlands, and Japan hold 80-95% market share in almost all SME segments, while China has about 2 percent market share across the same.[509]

**Market Share of Best Chinese SME Segments**[510]




In contrast, even China's SMIC relies on SME imports. Chinese news pieces, sometimes picked up by western reporting, have tended to vastly exaggerate Chinese SME capabilities.[511] In truth, China has almost no SME industry to speak of, including "only a small number of companies that are not at the state of the art."[512] Not even Chinese chipmakers buy Chinese SME, with less than 8 percent of China's annual SME demand fulfilled domestically in 2020, despite the goal set in "Made in China 2025" of reaching 50 percent localization.[513]

**SME: Lithography**

Of special importance within SME is lithography equipment, involved in transferring chip patterns onto raw wafers. Notably, as mentioned above, the Dutch company ASML is the only company in the world which produces extreme ultraviolet (EUV) lithography, necessary to produce all chips smaller than 7nm.[514]

The most advanced Chinese photolithography firm, Shanghai Micro Electronics Equipment (SMEE), on the other hand, tops out with only beginning to prototype the 90 nm node; that is, it is eight generations behind the leading node.[515] In turn, ASML is dependent on highly specialized components manufactured

[508] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 3-4; Khan, "Maintaining the AI Chip Competitive Advantage of the United States and its Allies," 4.

[509] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 11.

[510] Data source: Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 14.

[511] Ibid., 14.

[512] Khan and Mann, "AI Chips," 30.

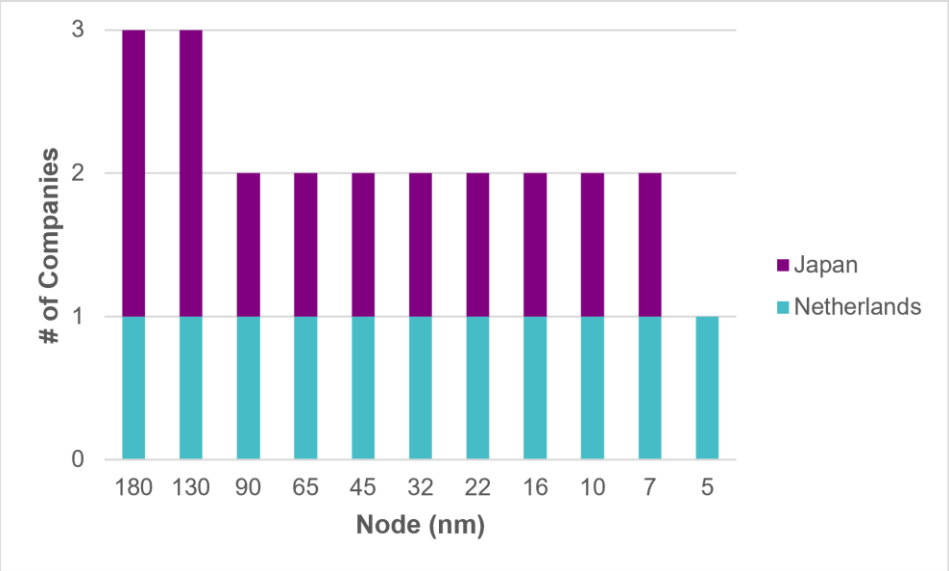
[513] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 11, 15.

[514] Kleinhans and Baisakova, "The global semiconductor value chain," 16-7.

[515] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 4.

by single firms in the US and allied countries, including special light sources from US-based Cymer.[516] Further, even the market for less advanced photolithography is highly consolidated – apart from ASML, the only other company producing photolithography equipment for chips past the 90nm node at scale is Japan-based Nikon.[517] Since the US has convinced the Netherlands to apply an export ban on EUV equipment to China, it seems very difficult for Chinese fabs to progress beyond the 7nm node "for the foreseeable future."[518]

**Photolithography Companies by Country and Node**[519]




**Other Materials**

Fabs also use various specialized chemicals and gases during chip manufacture, with strict impurity requirements of 1 part per billion or less. This segment is dominated by Japan's Shin-Etsu, Sumitomo Chemicals, and Mitsui Chemicals; and by Europe's BASF, Linde, and Merck KgaA.[520] Finally, fabs make chips out of wafers, a segment which has also rapidly consolidated. In 1990, more than 20 silicon wafer suppliers existed; in 2019, only 5 companies controlled 90% of the market: Japan-based Shin-Etsu and Sumco, Taiwan's GlobalWafers, Europe's Siltronic, and South Korea's SK Siltron.[521] In sum, across the many segments of the supply chain, China not only lacks indigenous capabilities, but in a significant number of cases is various generations behind the cutting-edge.

---

[516] Ibid.

[517] Khan and Flynn, "Maintaining China's Dependence on Democracies for Advanced Computer Chips," 4.

[518] Ibid., 8.

[519] Data source: VLSI Research.

[520] Kleinhans and Baisakova, "The global semiconductor value chain," 18.

[521] Ibid., 18.

### (c) What if China seizes Taiwan?

At this point, one might wonder: could China invade Taiwan to acquire control of TSMC?[522] As one industry analysis notes, "the importance of Taiwan for the semiconductor value chain cannot be overestimated. Almost the entire fabless industry for cutting-edge chips, globally, relies on TSMC – it is potentially the most critical single point of failure in the entire semiconductor value chain."[523] While this would not involve geographic dispersion of the ability to produce advanced AI chips, were Taiwan to be successfully absorbed by China, one could reasonably say that hardware production had then "diffused" there. However, even if China were willing to mount an invasion of Taiwan for chips, this seems unlikely to be succeed as a way of acquiring chip production capability.

First, TSMC would likely not remain operational following the invasion. Fabs comprise a large selection of delicate machines, with TSMC's fabs spanning cities across northern, southern, and central Taiwan. Hsinchu, home to TSMC's headquarters, also houses several Taiwanese military bases, sits on Taiwan's cost facing China, and hosts Taiwanese military drills practicing repelling a Chinese invasion.[524] Consequently, it seems likely that unless Taiwan were to surrender outright, most of the military contingencies currently planned for by both sides would carry a significant risk of destroying the machines it took TSMC many years to build.[525]

Second, even if Chinese forces sought to deliberately avoid damaging TSMC facilities, a difficult task given that the current dominant operational concept for Taiwan's defense involves long, protracted warfare on the island to impose costs on China and stall for US involvement, Taiwanese forces might destroy TSMC facilities themselves, to avoid having them fall into Chinese hands.[526] As World War II began, for example, the US government was not only able to persuade Standard-Vacuum, a US-based oil company in the Dutch East Indies, to participate in its oil embargo of Japan, but also secured an agreement from Standard-Vacuum that it would destroy its oil wells if the Japanese launched an invasion to seize its oil. Given that Taiwan would almost certainly depend on American intervention to survive a full-on Chinese invasion, in such a contingency its government would arguably see destroying TSMC – in the event that its facilities appeared about to fall into Chinese hands – as necessary to avoid strengthening the Chinese war effort against any incoming US forces.[527] Further, TSMC itself might also wish to destroy its Taiwan-based fabs to protect its own customer IP; being a global company with offices across Europe and North America, it might reason it possessed a greater chance of eking out survival if its technological secrets did not fall into Chinese hands.[528] And, of course, were the United States to actually engage in hostilities on Taiwan, it could easily destroy TSMC fabs itself, their composition representing simply a dozen or so fixed targets.

---

[522] Steve Blank, "The Chip Wars of the 21st Century," *War on the Rocks*, 2020, https://warontherocks.com/2020/06/the-chip-wars-of-the-21st-century/.

[523] Kleinhans and Baisakova, "The global semiconductor value chain," 22.

[524] "Taiwan military stages drill aimed at repelling China attack," *Associated Press*, January 18, 2021, https://abcnews.go.com/International/wireStory/taiwan-military-stages-drill-aimed-repelling-china-attack-75333944.

[525] Jon Stokes, "Why a Chinese invasion of Taiwan would be a catastrophe for China and the world," *Doxa*, 2021, https://doxa.substack.com/p/why-a-chinese-invasion-of-taiwan.

[526] Jim Thomas, John Stillion, and Iskander Rehman, "Hard ROC 2.0: Taiwan and Deterrence through Protraction," *CSBA*, 2014, https://csbaonline.org/uploads/documents/2014-10-01_CSBA-TaiwanReport-1.pdf.

[527] Jennifer Lind and Daryl G. Press, "Markets or Mercantilism? How China Secures Its Energy Supplies," *International Security* 42.4 (2018), 200.

[528] Stokes, "Why a Chinese invasion of Taiwan would be a catastrophe for China and the world."

Third, Chinese engineers might be unable to operate TSMC without its original employees, who may well have fled the country at the start of the conflict. As Stephen Brooks has argued, returns to conquest have tended to decline in the modern era in part because knowledge workers tend to be highly geographically mobile, and even when conquerors capture knowledge-based companies whole, it is easy for employees whose job it is to produce intangible outputs to engage in difficult-to-prevent shirking, compared to, say, the coal mines of yesteryear.[529]

Finally, perhaps most unavoidably, as discussed above, TSMC's chip manufacturing capabilities depend on raw inputs which are overwhelmingly sourced from American, European, and Japanese firms. In wartime, these inputs would almost certainly be cut off.[530] While existing stocks might provide a limited manufacturing capability, TSMC would not be able to continue operating for long.

### (2) Hardware Supply Cutoff Would Succeed

Even if convinced that China does not, and will not, for the near future, house the production of advanced AI chips, one might then ask: even if production is housed by the United States and its allies, would a supply cutoff succeed? After all, advanced AI chips are presently sold on the open market, and the history of export controls shows leaks are not exactly rare. Hugo Meijer argues, for example, that any export controls toward China will invariably fail, due both to lack of domestic political will and increased defection by allies.[531] Several reasons, however, point to likely success.

#### (a) Chokepoints have proliferated

First, the general tendency toward specialized, consolidated AI segments means a larger number of "bites at the apple" for effective cutoff. The United States itself has sole control over four chokepoints which would prevent China from producing leading-edge chips; even disregarding that, even if Japan decides to bandwagon with China, the US can still achieve a cutoff if Taiwan and South Korea agree to withhold their part of the supply chain, for example. Below, adapting work by Saif Khan, I show what chokepoints would be available if the US can secure a given number of allies' cooperation.

**Possible Chokepoints by Countries Needed**[532]

| US alone | US+2 | US+3 |
|---|---|---|
| EDA software | Advanced fabs (Taiwan, South Korea) | Leading-edge photomasks (Taiwan, Japan, South Korea) |
| x86 CPU design IP | EUV, ArFi photoresists (Japan, South Korea) | Photoresists (Taiwan, Japan, South Korea) |
| GPU design IP | ArFi scanners (Japan, Netherlands) | Advanced chemical vapor deposition (Japan, South Korea, Netherlands) |
| FPGA design IP | Photomask inspection, repair (Japan, Germany) | Logic chip test equipment (Taiwan, Japan, Italy) |
| **US+1** | E-beam lithography (Japan, Germany) | **US+4** |
| EUV, ArFi resist processing (Japan) | Laser lithography (Germany, Sweden) | Maskmaking tools (Japan, Germany, Sweden) |
| Wafer metrology, inspection (Japan) | Atomic layer etching (Japan, UK) | 300mm silicon wafers (Taiwan, Japan, South Korea, Germany) |
| Core chip IP (UK) | Advanced ion implanters (Taiwan, Japan) | Nanoimprint lithography (Japan, Germany, Sweden, Austria) |
| EUV scanners (Netherlands) | Rapid thermal processing (Japan, South Korea) | Wafer manufacturing equipment (Japan, Germany, Switzerland, Austria) |
| | Chemical mechanical planarization (Japan, South Korea) | |
| | Non-planar chip processes (Taiwan, South Korea) | |
| | AI ASIC design IP (UK, Israel) | |

[529] Stephen G. Brooks, *Producing Security: Multinational Corporations, Globalization, and the Changing Calculus of Conflict* (Princeton: Princeton University Press, 2006), 68-9.

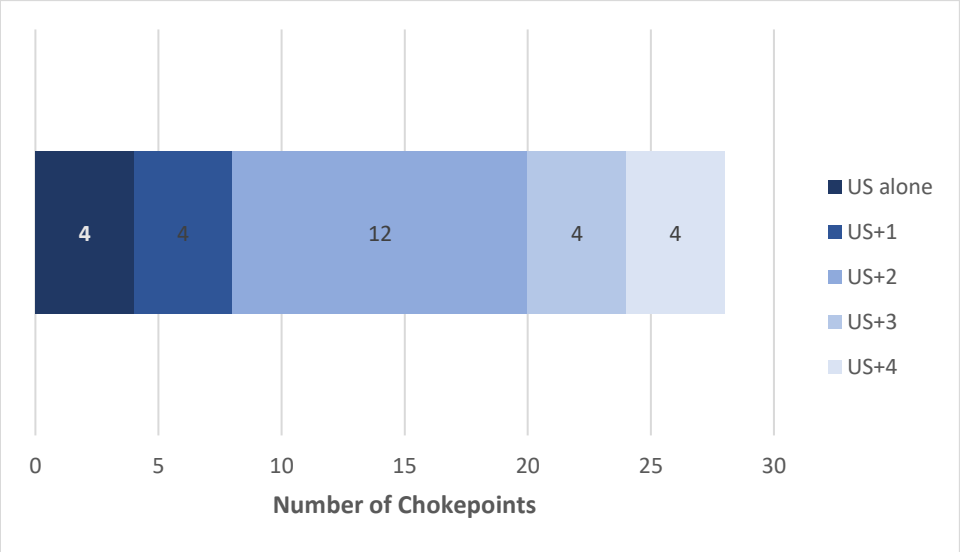
[530] John Lee and Jan-Peter Kleinhans, "Would China Invade Taiwan for TSMC?", *The Diplomat*, December 15, 2020, https://thediplomat.com/2020/12/would-china-invade-taiwan-for-tsmc/.

[531] Hugo Meijer, *Trading with the Enemy: The Making of US Export Control Policy Toward the People's Republic of China* (UK: Oxford University Press, 2016).

[532] Adapted from Khan, "Securing Semiconductor Supply Chains," 12-3, 41-2.

For legibility, we can display this visually: twelve more chokepoints are potentially available if the US can secure the cooperation of two additional countries, for example. Eugene Gholz argues that globalization makes countries more resilient to supply cutoff, because of the wide availability of alternative suppliers; here, however, the opposite seems true – for many segments, there are no alternative suppliers.[533]

**Chokepoints and Allies**




Importantly, chip design is fab-specific. That is, chip designs are highly complex blueprints which work with a specific fab; redesigning the same chip to be produced by a different fab can take years.[534] Consequently, were China to lose access to TSMC, even if a successful indigenous fab could be produced under desperate wartime measures (itself doubtful), a several-year pause in production would likely occur.

Further, even if the United States was unable to marshal allied willpower to cut off China, the fact of heavy US involvement in the AI supply chain also means ample opportunity for US exploitation of China-bound materials, potentially allowing the United States to decrease the effectiveness of Chinese AAWs even without denying China advanced AI chips altogether. According to 2021 Bloomberg reporting, for example, China-based supply chain elements inserted malicious chips into server motherboards made by US-based Super Micro Computer Inc. ("Supermicro"), which counts federal agencies among its clients. Consequently, DoD analysts worried that compromised Supermicro "implants" could "be a digital weapon that could shut down those systems during a conflict." Similarly, laptops from Lenovo Group Ltd., a US military supplier incorporated in Hong Kong, were found in 2008 to contain altered hardware which transmitted all inputted data back to China.[535] The United States presumably possesses hardware hacking capabilities that are at least as competent.

[533] Eugene Gholz and Daryl G. Press, "The effects of wars on neutral countries: Why it doesn't pay to preserve the peace," *Security Studies* 10.4 (2001): 1-57.

[534] Kleinhans and Baisakova, "The global semiconductor value chain,"" 14.

[535] Jordan Robertson and Michael Riley, "The Long Hack: How China Exploited a U.S. Tech Supplier," Bloomberg, February 12, 2021, https://www.bloomberg.com/features/2021-supermicro/?.

Finally, US ability to effectively impose supply cutoffs of relevant segments in coordination with allies, even in peacetime, has already been demonstrated. First, when China attempted to compete in DRAM by establishing the Fujian Jinhua Integrated Circuit Company in 2016, the US banned US-origin SME exports to that company in 2018, and it was forced to shutter all production after running out of critical US-dominated inputs in 2019.[536] When the US ban occurred, ASML voluntarily withdrew its engineers from the company on the same day. According to reporting by Reuters, Japanese SME firms also then decided to voluntarily mirror the US blacklist, declining to fill the market void left by the US withdrawal.[537]

Second, in December 2020, the United States imposed export controls against SMIC, issuing a presumption-of-denial policy banning export of items uniquely required for chip fabrication at 10nm and below.[538] Consequently, experts forecast that the fabrication market will not significantly change for at least ten years.[539] As mentioned, China has no companies capable of fabricating leading-node GPUs and FPGAs: Chinese chip design firms must turn to Taiwan's TSMC. While "Made in China 2025" set the goal of 70% self-sufficiency in semiconductors, local chip production remained at 16% in 2020, with market research firm IC Insights predicting that figure would only reach 19% in 2025.[540]

Finally, in 2019, the United States and the Netherlands began blocking Chinese access to critical SME, EUV in particular.[541] China only plans to develop an EUV machine by 2030; since China has already missed various other previously set goals about AI supply chain independence, even this is likely over-optimistic.[542] In December 2020, the US additionally increased the severity of its export controls on Huawei, preventing it from buying chips manufactured from fabs which use US SME. Huawei had stockpiled components for a full year amidst US trade tensions, but will at some point run out of up-to-date chips, and will likely be forced to pivot to other parts of its business.[543] Commenting on Chinese indigenization efforts responding to these developments, a manager at China's largest chip design toolmaker noted in 2021, "asking us to fully replace [US-based] Synopsys and Cadence is like coming to carmakers and asking to build rockets."[544]

Further controls are also possible. China is so enormously behind in this segment that more complete SME controls would likely block Chinese chip fabrication beyond the 90nm node, eight generations behind

[536] Matthew Broersma, "Chinese Chip Giant Fujian Jinhua To Cease Operations After US Ban," Silicon, January 2019, https://www.silicon.co.uk/workspace/chinese-chip-giant-fujian-jinhua-to-cease-dram-production-after-us-ban-240903.

[537] Khan, "U.S. Semiconductor Exports to China: Current Policies and Trends," 12-3.

[538] Harry Clark, W. Clark McFadden II, and Jeanine McGuinness, "The New 'Uniquely Required' Standard of U.S. Export Licensing Policy for SMIC," JD Supra, 2021, https://www.jdsupra.com/legalnews/the-new-uniquely-required-standard-of-u-2185333/.

[539] Kleinhans and Baisakova, "The global semiconductor value chain," 16.

[540] Cheng Ting-Fang and Lauly Li, "US-China tech war: Beijing's secret chipmaking champions," *Nikkei Asia*, 2021, https://asia.nikkei.com/Spotlight/The-Big-Story/US-China-tech-war-Beijing-s-secret-chipmaking-champions.

[541] Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 8.

[542] Ibid., 8, 15.

[543] Khan, "U.S. Semiconductor Exports to China: Current Policies and Trends," 12-3; Kleinhans and Baisakova, "The global semiconductor value chain," 13.

[544] Ting-Fang and Li, "US-China tech war."

current leading-edge 5nm production.[545] According to one market model, even controlling SME at or past 45nm would cut China's share of global chip fab capacity all the way to 0.2%.[546]

**(b) What about stockpiling?**

One might ask: since great powers are likely to stockpile a significant inventory of weapons, will supply cutoffs will lack sufficiently immediate effects to deter wars?[547] However, this argument is unlikely to apply to AI. First, since leading-node chips are extremely finite, especially at first, states are unlikely to be able to build up very large stockpiles of AAWs. Second, as mentioned above, in the case of Taiwan contingencies, the island government is likely to attempt to stall, in the event of Chinese invasion, for US arrival. This may lead to a naturally protracted conflict.[548] Third, this stockpile would repeatedly obsolete itself as chips changed nodes.

Finally, most importantly, the consensus view of scholarship is that AAWs are likely to exactly encourage wars of attrition, as the relative lack of human casualties will reduce the probability that either side exits due to public outcry. Thus, possessing a lesser stockpile of AAWs will likely be especially relevant if predicted changes to warfare come to pass. [549]

**(c) Could China cannibalize its civilian sector?**

Following a supply cutoff, could China simply cannibalize civilian-sector AI chips for military use? After all, advanced AI chips are a ubiquitous consumer goods input, including into iPhones, laptops, and cars. Analogously, when considering the effects of a blockade of oil on China, analysts have found that commandeering domestic supplies could help cushion against a distant US supply cutoff of the Strait of Malacca, potentially allowing remaining supply lines (e.g., overland pipelines) to suffice for China's military needs.[550]

With AI chips, however, this is unlikely. First, unlike with oil, there are presently no countries which are both geopolitically aligned with China and possess leading-node capabilities. Russia and Central Asia are players in oil and natural gas and so can relieve China by land, avoiding US maritime dominance, but no equivalent actors exist for AI chips. Second, leading-node chips used in consumer goods are not precisely fungible with those likely to be used in AAWs. Chips are designed to be produced by specific fabs, and customized for specific purposes. Consequently, one cannot, say, order chips for laptops, but then reroute those chips for AAWs. It would likely be necessary to design the chips for AAWs in the first place.

**(d) What about transshipment?**

Finally, the reader may ask, what about third parties making purchases on behalf of China? While transshipment of oil has frequently been used to avoid sanctions, this seems less likely in the case of advanced AI chips. First, most fundamentally, as mentioned above, chips are not wholly fungible. Second,

---

[545] Khan, "Maintaining the AI Chip Competitive Advantage of the United States and its Allies," 5.

[546] Khan and Flynn, "Maintaining China's Dependence on Democracies for Advanced Computer Chips," 8.

[547] Jonathan Kirshner, "The Changing Calculus of Conflict?", *Security Studies* 16.4 (2007): 583-597.

[548] Thomas, Stillion, and Rehman, "Hard ROC 2.0."

[549] Zegart, "Cheap fights, credible threats"; Gartzke, "Blood and robots"; Lin-Greenberg, "Remote Controlled Restraint."

[550] Gabriel B. Collins and William S. Murray, "No Oil for the Lamps of China?", Naval War College Review 61.2 (2008): 1-17. Available online: https://digital-commons.usnwc.edu/nwc-review/vol61/iss2/10/.

even were this possible, the world is consuming advanced chips, an extremely finite resource, faster than they are being produced. Consequently, even when TSMC cut off supplying Huawei due to US policy change under the Trump administration, it suffered no net loss of business – in the words of CEO Peter Wennink of ASML, the European company with, again, a complete monopoly on EUV equipment required to produce all leading-edge chips, "if we cannot ship to customer A or country B, we'll ship it to customer C and country D."[551] Unlike oil, a difficult-to-control commodity sometimes described as "one big pool," all of the world's advanced chips are spoken for and highly valued by specific clients. As a result, especially under wartime conditions, surveilling and protecting the AI supply chain would likely be a top priority for the United States and its allies.

## Conclusion

Regardless of whether fully autonomous drones revolutionize future warfare, control over production of advanced AI chips by the United States and allies almost certainly means such drones would not, if they emerge, advantage China. In 2019, China imported $238 billion of crude oil, but $301 billion of semiconductors; a supply cutoff is both feasible and could attack any number of parts of the supply chain.[552]

Although China is engaged in intense indigenization efforts, a consensus forecast is that achieving chip self-sufficiency is extremely unlikely within the next 10 years.[553] At a minimum, a very conservative estimate is that it is very difficult to imagine that China will successfully develop EUV technology, a key bottleneck for the most advanced chips, for at least that timeframe.[554] Less conservatively, without US, Japanese, and Dutch components, one Brookings report assessed that China "would find it nearly impossible to develop or maintain advanced chip factories for the foreseeable future."[555]

Consequently, if AAWs are realized within the next decade, the US will likely enjoy uncontested use for a notable period of time. Contrary to analysts predicting China will leapfrog the United States using AAWs, the opposite may be true – insofar as other trendlines look unfavorable for the United States in the conventional domain vis-à-vis China, the rise of AAWs as a significant military technology could instead restore US advantage in the Western Pacific balance.

Besides commenting on the specific question of AAWs and the US-China balance, this paper arguably makes several additional theoretical contributions. First, these results advance the general study of AI by political science. Existing assessments suffer from scarce data – AI-enabled weapons have not yet been widely fielded *en masse* or used in wars, precluding empirically-driven analysis. To compensate, this paper

[551] Khan and Flynn, "Maintaining China's Dependence on Democracies for Advanced Computer Chips," 9.

[552] Kleinhans and Baisakova, "The global semiconductor value chain," 5.

[553] Shannon Davis, "China to Fall Far Short of its 'Made-in-China 2025' Goal for IC Devices," Semiconductor Digest, May 21, 2020, https://www.semiconductor-digest.com/2020/05/21/china-to-fall-far-short-of-its-made-in-china-2025-goal-for-ic-devices/. Similarly, according to a survey of experts carried out by the Center for Security and Emerging Technology, China will likely take at least ten years to catch up in the photolithography, deposition, and etch SME segments. Of course, even if this occurs, the US will still hold various other chokepoints. See Hunt et al., "China's Progress in Semiconductor Manufacturing Equipment," 16.

[554] Saif M. Khan, "Maintaining the AI Chip Competitive Advantage of the United States and its Allies," CSET, 2019, https://cset.georgetown.edu/wp-content/uploads/CSET-Maintaining-the-AI-Chip-Competitive-Advantage-of-the-United-States-and-its-Allies-20191206.pdf, 4.

[555] Khan and Flynn, "Maintaining China's Dependence on Democracies for Advanced Computer Chips," 1.

focuses on *inputs* instead of *outputs*, illustrating how existing datasets can be brought to bear – analogously, demographics permits estimating how large the military-age population for some state may be, several decades out, even if that state's future military doctrine is harder to predict.

Second, the paper intervenes in the related literature of whether liberalism has superseded mercantilism. This ongoing debate asks whether global markets have dampened causes of conflict related to scarcity, and centers on oil as its primary case. One side argues that commoditization means attempts at coercion through resource denial translate only into relative price changes (as with the diversification of upstream crude oil producers, after the 2000s), rather than existential threats (as with the US embargo on Japan which helped motivate Pearl Harbor).[556] For Eugene Gholz, for example, threats of supply cutoffs are a "weak reed," as globalization has meant an increase in alternative sources of supply – the US embargo on Japan was only workable in a previous era.[557] Others argue even globally traded goods contain actors with disproportionate influence, with interdependence being variously weaponizable by central nodes.[558] Here, Gholz replies by suggesting that defense industries are especially unlikely to "rely on links for which substitutes are unavailable," but in the case of the AI supply chain, industries have no choice when oligopolies or even monopolies obtain in particular segments.[559] This paper's analysis of the AI supply chain provides evidence that even as interdependence can generate economic incentives for peace, goods with certain technical characteristics can give rise to specific, highly concentrated market segments which then represent sources of coercive power.[560]

Finally, the paper makes several contributions to the diffusion literature. AI as a case usefully tests competing mechanisms in the extant literature predicting diffusion, or its absence: AI is an unusually open

---

[556] Robert J. Weiner, "Is the World Oil Market 'One Great Pool?'" *Energy Journal* 12.3 (1991), 95-108; Rosemary A. Kelanic, "Black Gold and Blackmail: The Politics of International Oil Coercion" (PhD diss., University of Chicago, 2012); Llewelyn Hughes and Phillip Y. Lipscy, "The Politics of Energy," *Annual Review of Political Science* 16.1 (2013), 449–69; Michael Levi, "The Enduring Vulnerabilities of Oil Markets," *Security Studies* 22.1 (2013), 132–138; Eugene Gholz and Daryl G. Press, "Enduring Resilience: How Oil Markets Handle Disruptions," *Security Studies* 22.1 (2013), 139-47l; Gabriel B. Collins and William S. Murray, "No Oil for the Lamps of China?," *Naval War College Review* 61.2 (2008), https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1913&context=nwc-review; Michael Mousseau, "The End of War: How a Robust Marketplace and Liberal Hegemony Are Leading to Perpetual World Peace," *International Security* 44.1 (2019): 160-96; Brooks, *Producing Security*; Danielle F.S. Cohen and Jonathan Kirshner, "The Cult of Energy Insecurity and Great Power Rivalry Across the Pacific," in *The Nexus of Economics, Security, and International Relations in East Asia*, eds. Avery Goldstein and Edward D. Mansfield (US: Stanford University Press, 2012), 144-76.

[557] Eugene Gholz, "Globalization, Systems Integration, and the Future of Great Power War," Security Studies 16.4 (2007): 615-36. Also Kirshner, "The Changing Calculus of Conflict?", 635.

[558] Jennifer Lind and Daryl G. Press, "Markets or Mercantilism? How China Secures Its Energy Supplies," *International Security* 42.4 (2018): 170-204; Henry Farrell and Abraham L. Newman, "Weaponized Interdependence: How Global Economic Networks Shape State Coercion," *International Security* 44.1 (2019): 42-79; Daniel W. Drezner, Henry Farrell, and Abraham L. Newman, eds., *The Uses and Abuses of Weaponized Interdependence* (Brookings, 2021); Llewelyn Hughes and Austin Long, "Is There an Oil Weapon? Security Implications of Changes in the Structure of the International Oil Market," *International Security* 39.3 (2014/15): 152-89; Jonathan D. Caverley, "United States Hegemony and the New Economics of Defense," *Security Studies*, 16.4 (2007), 598–614.

[559] Ibid., 619.

[560] Lind and Press, "Markets or Mercantilism?", 200; John Lee and Jan-Peter Kleinhans, "Would China Invade Taiwan for TSMC?", *The Diplomat*, December 15, 2020, https://thediplomat.com/2020/12/would-china-invade-taiwan-for-tsmc/.

field with strong commercial applications, even among modern information technologies, but is also a complex, investment-intensive technology commonly understood as important to future military power. Consequently, observing how clashing causal mechanisms interact in the AI case naturally horse-races extant theories of diffusion against each other.[561]

Further, the paper arguably illustrates the utility of third-image, system-level analyses of the diffusion of a particular technology, in addition to predominant second-image, state-specific explanations which focus on states' finances, internal politics, organizational capital, and culture.[562] Extant second-image narratives adopt this basic framing: states wish to adopt the best technologies available, and so the question of diffusion relates to which states are able to do that, as determined by differently emphasized material or intangible factors.[563] Consequently, recent discussion over whether globalization has accelerated or inhibited diffusion centers on whether more or less is now demanded of would-be adopting states. Does the internet, increased trade, and the ease of transmission and theft of assets like software reduce adopting states' burdens, or do specialization, increased complexity, and the shift from codified to tacit knowledge imply the opposite?[564] This paper suggests that diffusion patterns are controlled by system-level variables, however, in addition to state-specific factors. Theoretically, the argument outlined above demonstrates that for some innovations, strategic, economic, and/or scientific limits mean diffusion will be limited to some fixed quantity of states, regardless of the absorptive capacity of any would-be adoptee in question.[565]

Finally, the extant diffusion literature importantly fails to distinguish between ***production*** and ***acquisition.*** Whether or not states produce technologies domestically or acquire them abroad is infrequently discussed. Instead, adoption is typically operationalized as binary: either a state adopts a given innovation or does not.

---

[561] Andrea Gilli and Mauro Gilli, "Why China Has Not Caught Up Yet: Military-Technological Superiority and the Limits of Imitation, Reverse Engineering, and Cyber Espionage," *International Security* 43.3 (2018/19): 141-89; Michael Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1 (2018), https://doi.org/10.15781/T2639KP49; Hugo Meijer, *Trading with the Enemy: The Making of US Export Control Policy Toward the People's Republic of China* (UK: Oxford University Press, 2016); Richard A. Bitzinger, "The Globalization of the Arms Industry: The Next Proliferation Challenge," *International Security* 19.2 (1994), 170-198; Joseph S. Nye Jr., *The Future of Power* (New York: PublicAffairs, 2011); Stephen G. Brooks, *Producing Security: Multinational Corporations, Globalization, and the Changing Calculus of Conflict* (Princeton: Princeton University Press, 2006).

[562] Michael C. Horowitz, *The Diffusion of Military Power: Causes and Consequences for International Politics* (Princeton: Princeton University Press, 2010); Emily O. Goldman and Leslie C. Eliason, eds., *The Diffusion of Military Technology and Ideas* (California: Stanford University Press, 2003); Emily O. Goldman and Richard B. Andres, "Systemic effects of military innovation and diffusion," *Security Studies* 8 (1999): 79-125; William C. Wohlforth, Dmitry Adamsky, Theo Farrell, Adam Grissom, Thomas G. Mahnken, and Michael C. Horowitz, "H-Diplo/ISSF Roundtable Review of Michael C. Horowitz. The Diffusion of Military Power: Causes and Consequences for International Relations (2010)," *Roundtable* 3.10 (2012), https://issforum.org/ISSF/PDF/ISSF-Roundtable-3-10.pdf.

[563] Waltz, *Theory of International Politics*, 127; Joao Resende-Santos, *Neorealism, States, and the Modern Mass Army* (New York: Cambridge University Press, 2007); Emily O. Goldman, "Cultural foundations of military diffusion," *Review of International Studies* 32.1: 69-91; Kimberly Zisk, *Engaging the Enemy: Organization Theory and Soviet Military Innovation, 1955-1991* (Princeton: Princeton University Press, 1993), and others.

[564] Gilli and Gilli, "Why China Has Not Caught Up Yet"; Meijer, *Trading with the Enemy;* Brooks, *Producing Security*.

[565] Earlier theorizing about nuclear weapons took a similar tack, arguing that achieving a monopoly might permit the the first-mover to preclude adoption by any other state. The best history of this discussion is George H. Quester, *Nuclear Monopoly* (UK: Routledge, 2000).

For example, under Horowitz's "adoption-capacity theory," states either pursue adoption, or select other responses depending on the innovation's financial intensity and organizational capital requirements; Horowitz further decomposes financial intensity into two factors: whether the "underlying basis of the technology" is civilian or military, and the cost per unit of the technology.[566] Similarly, Daniel Drezner's 2x2 typologizes technologies using two factors: whether the private or public sector leads, and whether the fixed-cost investments needed "to develop or adopt" the new technology are high or low.[567] In a collection of essays about technological diffusion, Emily Goldman and Leslie Eliason describe diffusion simply as occurring when "an idea, thing, or practice is transmitted from one social group to another."[568]

In significant part, inattention to production has been the norm because studies of diffusion have sought to produce theory applicable to a very broad range of phenomena, both material and immaterial: that is, discussions of diffusion freely range not only over weaponry requiring new complex, physical materials, such as tanks or aircraft carriers, but also over intangibles, such as ideas and doctrines whose utility emerges primarily from making use of *existing* materials (e.g., suicide terrorism), and which therefore do not acquire production to adopt.[569] While this is understandable, it also means that significant swaths of theorizing in the diffusion literature have somewhat lost touch with the actual physicality of military weapons. Illustratively, Horowitz's nuclear weapons case study focuses entirely on financial costs as the key limitation for would-be proliferators; he does not discuss the limited availability of fissile material in the early nuclear era. This is an important omission, I claim, because some things are not for sale – the United States would not have sold off its nuclear weapons for any amount of money, in 1955.[570]

In other words, even if a given technology "diffuses" in the sense of being possessed in some quantity by, and competently used by, various militaries, this is not equivalent to the diffusion of all of the power which flows from that technology. AAWs may become widely used – in fact, the US military may sell them abroad itself, and China may even get its hands on appreciable quantities. This is not, however, equivalent to diffusion: in reviewing the almost universal failure of academic theories to predict North Korea's successful nuclearization, for example, Nicholas Miller and Vipin Narang note the importance of the eroding effectiveness of supply-side control measures, given that nuclear technology is now decades old and counting.[571] Since North Korea indigenously and secretly developed its nuclear reactor and reprocessing facility, its reduced elicitation of sensitive nuclear assistance from other countries helped it fly under the

[566] Horowitz, *The Diffusion of Military Power*, 24-32.

[567] Daniel W. Drezner, "Technological change and international relations," *International Relations* 33.2 (2019), 292.

[568] Emily O. Goldman and Leslie C. Eliason, "Introduction," in *The Diffusion of Military Technology and Ideas* (California: Stanford University Press, 2003), eds. Emily O. Goldman and Leslie C. Eliason, 11.

[569] Suicide terrorism is Horowitz's fourth case study. See *The Diffusion of Military Power*, 166-207.

[570] Ibid., 98-133.

[571] Nicholas L. Miller and Vipin Narang, "North Korea Defied the Theoretical Odds: What Can We Learn from its Successful Nuclearization?," *Texas National Security* Review 1.2 (2018), https://tnsr.org/2018/02/north-korea-defied-theoretical-odds-can-learn-successful-nuclearization. See also Francis J. Gavin, "Strategies of Inhibition: U.S. Grand Strategy, the Nuclear Revolution, and Nonproliferation," *International Security* 44.2 (2019), 185-92; and Nicholas L. Miller, *Stopping the Bomb: The Sources of Effectiveness of US Nonproliferation Policy* (US: Cornell University Press, 2018); R. Scott Kemp, "The Nonproliferation Emperor Has No Clothes: The Gas Centrifuge, Supply-Side Controls, and the Future of Nuclear Proliferation," *International Security* 38.4 (2014): 39-78.

radar of detection by the United States.[572] In other words, nuclear weapons diffused to North Korea because of a change in the underlying difficulty of production. Those studying AI would do well to pay attention to the comparable underlying dynamics of advanced chips.

Thus, the analogy to uranium may be suitable in at least one more way. At the open of the nuclear age, scholars generally predicted rapid proliferation: since nuclear weapons provided their owners with such enormous military advantages, it was thought, their acquisition would be widely desired. Decades later, however, nuclear-armed states remain in the single digits. Presently, a growing literature debates whether AI-enabled weapons of the future will revolutionize warfare, and hence proliferate widely out of state desire to keep up with the times, or whether another AI winter will set in, following eventual disappointment with AI's inevitable limitations. Thinking of advanced AI chips as fissile material, however, opens another possibility – AAWs may come to both be regarded as revolutionary, but also fail to diffuse widely in their first decades, due not to financial intensity, demands upon organizational capital, or complexity, but due to careful control by the United States and its allies of the key intermediate precursors to such weapons.

In the long-run, of course, even if previous-generation AAWs cannot defeat those fielded by the United States, they may still be used in various regional conflicts, or by states eager to police their own populations. I also do not claim this imbalance in advanced chip access is irrevocably permanent. It is not maintained by physics, after all, but merely a present confluence of economics and geopolitics.[573] Such imbalances, especially when used in war, tend toward their own erasure – after the 1973 oil crisis, when the Organization of Arab Petroleum Exporting Countries (OAPEC) embargoed Israel-supporting nations during the Yom Kippur War, the United States undertook intense efforts to develop alternative energy sources.[574] If China were to suffer defeat in conventional conflict with the United States due to lacking access to AAW precursors, and if the rivalry were to persist following that conflict's political consequences, avoiding a repeat scenario would likely become a top priority for the following decades. Consequently, I limit the scope of the paper's argument to ten years.[575]

[572] Miller and Narang, "North Korea Defied the Theoretical Odds." On sensitive nuclear assistance, see Matthew Kroenig, Exporting the Bomb: Technology Transfer and the Spread of Nuclear Weapons (Ithaca: Cornell University Press, 2010).

[573] As R. Scott Kemp wrote about the increasing impracticality of denying countries with nuclear ambitions access to gas centrifuges, "What was once exotic is now pedestrian, and nuclear weapons are no exception." Kemp, "The Nonproliferation Emperor Has No Clothes," 41.

[574] According to declassified British documents, the United States may have considered responding by invading Saudi Arabia, Kuwait, and Abu Dhabi with two or three brigades, followed by two divisions, to seize its oilfields. Michael Peck, "The Time America Almost Invaded OPEC," *The National Interest*, 2014, https://nationalinterest.org/feature/the-time-america-almost-invaded-opec-15726.

[575] Of course, looking more than ten years out, it also becomes increasingly difficult to predict whether current geopolitical conditions will continue. For example, China may have democratized, or suffered economic collapse; conventional conflict between the United States and China may have become increasingly unthinkable for any number of reasons.

# Conclusion

This dissertation defends an operationalization of artificial intelligence as deep learning, then assesses its domain-by-domain impact on the US-China balance. In doing so, it seeks to offer AI to security studies as a legible object, breaking ground on efforts to apply known methods to study an emerging technology. In the nuclear domain, it finds that AI could counterfactually enable US counterforce against China. In the conventional domain, it finds that the AI hardware supply chain strongly favors the United States in any contest of AAWs which unfolds in at least the next ten years.

## Avenues for Future Research

This dissertation suggests several avenues for future work. We can usefully divide such avenues into projects and methods, and consider them in turn.

## Projects

While AI has received rapidly increasing attention from economics, sociology, and philosophy, including its self-spawned fields of AI ethics and AI governance, to say nothing of computer science itself, it has curiously been the subject of relatively scant attention from political science and security studies. For this dissertation, the bulk of cited literature has come not from academia proper, but from outlets contained by, or adjacent to, the national security ecosystems of the world. In my view, this is regrettable.

In the nuclear domain, the campaign analysis undertaken in the second paper could be extended to the US-China undersea balance, as well as to other dyads. It will also be important to consider under what circumstances, or given what additional enabling technologies, AI may mean purely conventional counterforce is possible, both for the United States against China, as well as elsewhere.

The AI applications discussed as changing the US-China nuclear balance should also change the conventional balance, both in that dyad and more generally. Given the proliferation of precision-strike capabilities, advantaging finders over hiders is very likely useful not only for locating TELs and submarines, but also for defeating concealment in purely conventional conflicts. For example, given its ability to rapidly process vast quantities of data, deep learning will also alter global naval balances, advantaging those with access to SAR constellations by easing the search of oceans for surface vessels.[576]

Outside security studies, a precise study of how AI affects the political economy of authoritarianism ought to be conducted. To what degree will AI shift the incentives of elites? Full answers should emerge from comparative politics and dissection of the history of previous potentially centralizing technologies. For example, could AI reduce the size of the minimum winning coalition within political systems by automating many of the instruments of power?[577] Could AI dampen the negative economic consequences of internal surveillance? Are automated armies a solution to coup-proofing?

## Methods

First, I enumerate several methodological approaches at the end of this dissertation's third essay, including application of existing IR theory, mass and expert surveys, wargames, formal modeling, and campaign analyses. The second essay, of course, serves as an example of a campaign analysis; the third essay, in turn, exemplifies the use of data about *inputs*, even as the battlefield outputs of key AI resources remain immature

[576] C.P. Schwegmann, W. Kleynhans, B.P. Salmon, L.W. Mdakane, and R.G.V. Meyer, "Very deep learning for ship discrimination in Synthetic Aperture Radar imagery," paper presented at the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), China, 2016, https://ieeexplore.ieee.org/abstract/document/7729017/.

[577] Bruce Bueno de Mesquita, Alastair Smith, Randolph M. Siverson, and James D. Morrow, *The Logic of Political Survival* (US: MIT Press, 2003).

and difficult to study. This focus on inputs could be variously extended. The most low-hanging fruit, for one, is that at time of writing, revenue data about the ASICs discussed in this dissertation's third essay, unlike GPUs and FPGAs, appears unobtainable, due to technological novelty. Since whether or not AI hardware diffuses will importantly determine the degree to which China (and other countries) can leverage AAWs (and other AI-enabled military applications), it seems easy to say that someone ought to write, in a few years, an analysis with all available data at that time of to what degree that diffusion has occurred, especially in light of the recent US rediscovery of various forms of industrial policy (including likely large subsidies of various US companies involved in the AI supply chain, as well as in convincing TSMC to invest in some amount of capability in the United States). As technological progress continues to clarify what forms of data and engineering talent are most critical for battlefield applications, similar analyses could be done for data and AI talent, as well.

Second, albeit somewhat farther afield, I have come to believe political scientists ought to consider significant forays into forecasting, both of technological futures and of political outcomes more broadly, as a partial solution to the problems of data scarcity in studying actively unfolding phenomena. Most works in political science are already at least predictive – for example, campaign analyses can be understood as making forecasts about the likely outcomes of military contingencies, subject to specified scope conditions. More fundamentally, any theoretical partition of the possible values of some variable represents, arguably, a forecast that those possibilities should manifest more than an infinitesimal fraction of the time. On one end, grand theory is often too loose to generate point predictions useful to policymakers, and among scholars this tends to provoke decades of essentially unresolvable debates between theoretical schools. On the other end, quantitative hypothesis-testing on datasets of definitionally past occurrences has tended to encourage a theory-free search for empirical regularities, but it is often unclear whether such regularities map usefully onto the future.[578] While expert predictions have an abysmal track record, evidence suggests the accuracy of such predictions can be improved with a little training and effort, especially for nearer-term forecasts.[579] Since policymaking must deal with ongoing events, a willingness to consider explicit near-term forecasts based on one's work would naturally increase its real-world relevance; conversely, generating predictions should also avoid endless theoretical speculation in a vacuum – if one's work generates no testable propositions, after all, then it is definitionally non-falsifiable. With AI, the very broad range of implicit forecasts made about its impact on international affairs makes some kind of forecasting easy in applying political science research, but this idea could be extended to other subjects, as well.[580]

## Implications for Scholars

This dissertation offers several contributions to security studies. First, for readers convinced of AI's importance, the first essay offers a focus on deep learning as a productive direction for the field. The second and third essay extend this focus to the nuclear and conventional domains, illustrating how a focus on specific applications informed by technical detail can generate epistemic progress even though AI capabilities have not yet been deployed at scale on the world's battlefields.

---

[578] John J. Mearsheimer and Stephen M. Walt, "Leaving theory behind: Why simplistic hypothesis testing is bad for International Relations," *European Journal of International Relations* 19.3 (2013): 427-57. See also the discussion of difficulties with pooling data in Gary King, "Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data," *International Organization* 55 (2001): 497-507.

[579] Tetlock, *Expert Political Judgment*; Philip E. Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction* (NY: Broadway Books, 2016); J. Peter Scoblic and Philip E. Tetlock, "A Better Crystal Ball: The Right Way to Think About the Future," *Foreign Affairs*, 2020, https://www.foreignaffairs.com/articles/united-states/2020-10-13/better-crystal-ball.

[580] Existing forecasting efforts which could be interesting sources of data for political scientists include Metaculus (https://www.metaculus.com/), Foretell (https://www.cset-foretell.com/), and of course the Good Judgment Project (https://goodjudgment.com/).

Second, for those who study nuclear weapons, this dissertation offers evidence for simulation that AI may undermine mobility as a strategy for arsenal survivability, demanding new countermeasures from states previously reliant on road-mobile missiles to guarantee their second-strike capability. Of course, this effect will apply unevenly across states, but even for China, many of the available countermeasures may increase escalation risks, as they involve either destructive interference with US nuclear systems themselves, or progressive movement toward postures more toward the "always" branch of the always/never dilemma. To stop enemy AI from converting their mobile platforms into, essentially, large fixed targets, states will have to choose from various possible countermeasures outlined in the third essay. Worse-off states, like Pakistan, Iran, and North Korea, may have difficulty in the near to medium term in fully restoring their survivability.

Finally, for theorists of military power, this dissertation shows that technological progress may produce highly concentrated supply chokepoints which favor dominant states, rather than diffuse power to the periphery through globalized interdependence. Advanced AI chips are firmly ensconced on US and allied soil, with transnational production having led to monopolistic specialization rather than the easy redundancy of oil-like commodities available from many suppliers. If US supply cutoffs would be effective, as the third essay argues they would be, then technological innovation at the increasingly impenetrable R&D frontier may re-concentrate power in the United States, rather than promote its gradual erosion.

## Implications for Policymakers

Finally, this work also has implications for policymakers. First, the sky is not falling – if China displaces the United States as the world's leading power in the near future, AI will not be the cause. The geography of AI hardware production strongly favors the United States in any duel over the conventional applications of AI for at least the next ten years, and at higher levels of escalation, AI may even enable successful US counterforce against the Chinese nuclear arsenal. Consequently, risks to the United States in adopting AI technologies emerge more from unsafe implementations borne out of corner-cutting and unnecessary haste, rather than from falling behind. While a healthy sense of competition is indispensable, running so fast while ahead that we trip and smash our collective face into the ground would be senseless.

Second, China may feel increasingly threatened by US counterforce ability over the coming decade. AI-enabled intelligence-processing means the US ability to hold the Chinese arsenal at risk will increase naturally, even without the specific acquisition of further counterforce-specific platforms. US policymakers should decide whether they endorse such a development, and consciously choose whether to proceed; strategy should dictate technology, not vice versa. Bilateral and trilateral discussions with China and/or Russia about these issues should follow that strategic determination. Further, American foreign policy in regions with other competitive nuclear dyads, such as India and Pakistan, must also attend to the increasing possibility of counterforce by American allies, including the attendant escalation risk.

Finally, this work offers several guidelines for best competing with China in AI. Basic, private-sector AI research will almost certainly all diffuse, but US AI hardware and talent advantages are worth investing in. Strong technical barriers, such as those surrounding EUV photolithography, mean China will be unable to produce its own advanced AI chips for at least the next ten years, a horizon which prudent policy could extend. Conversely, the United States should make sure that its own government-sponsored research does not further some of AI's more authoritarian possibilities, including those now sadly on display in Xinjiang. Those AI fields, such as research underpinning civilian surveillance technologies, disproportionately further Chinese strategic ends. As the country leading the world in AI research, the United States can shape the direction of the technology to its strategic advantage. If those forecasting civilization-changing effects for more advanced AI are anywhere close to correct, the future of US prosperity may depend on it.

# Appendix – Simulation Code

The second essay centers on simulating US nuclear counterforce efforts against China. We intend to continue iterating on the code, and will host the latest version on Github.[581] Below, we document the function of each code module at time of writing to make the structure of our program more legible.

**config.py**
Stores all parameters used by the model, including quantitative assumptions about TELs, US intelligence capabilities and nuclear arsenal, and China itself.

Configuration is stored in a series of Python classes that support overriding, so that the default config can be modified to create specialized configurations to represent different Chinese strategies under low, high, and medium alert levels.

**simulation.py**
Contains the Simulation object which acts as a container for all other objects, and performs basic functions of running the simulation, including:

- Keeping track of the current time. Time in the simulation begins at noon on January 20th, 2021, and progresses in steps of one minute. Real times are used rather than offsets from an arbitrary "t=0" so that sunrise and sunset times can be simulated realistically.
- Running the simulation's event loop.

The simulation supports two main modes of operation:

- In Base Local mode, each TEL is considered to stay within a radius of its home base, returning after each trip. In this mode, TELs are linked together based on base affiliation, so e.g. TELs associated with the same base are assumed to experience the same weather.
- In Free Roaming mode, each TEL is assumed to wander throughout all of China independently, increasing the amount of sensor data that must be processed in order to find all of the TELs.

**enums.py**
Contains low level data types (enumerations) used throughout the simulation, representing concepts such as the different states a TEL can be in, the different detection methods available to the US, and so on.

**location.py**
Contains a Location class, representing a (lat, lon) coordinate pair. The main use of the Location object in this simulation is to allow calculating sunrise and sunset times for a given TEL base (determining whether nearby TELs are visible to EO satellites). This calculation takes into account both geographical position and the time of year.

**tel.py**
Contains the TEL class, which represents either a TEL, or a Chinese decoy which is designed to look and behave similarly to a TEL. Each TEL independently transitions through a configured set of states based on the alert level (e.g. staying in base for 16 hours, then roaming for 8 hours). In addition, while in Free Roaming mode each TEL independently simulates weather conditions.

The TEL class also contains methods for reporting how far the TEL has roamed since it was last observed, and how many square km must be destroyed in order to cover all of the areas it could have roamed to in this time.

[581] https://github.com/torinmr/nuclear-simulation.

**tel_base.py**
Contains the TELBase class, which holds a collection of TELs based at the same location. In Base Local mode, it also simulates local weather.

This module also contains functions for loading the set of TELs from an external data file.

**intelligence_types.py**
Contains data types used to represent different stages of US collection and processing of intelligence data.

**TLO:** A TEL-Like-Object. Represents an object which the US may (rightly or wrongly) believe is a TEL. Every TEL and decoy has a corresponding TLO, and there are also TLOs to represent heavy trucks which could be mistaken for TELs. A single TLO object can represent multiple entities in the simulation, so it is possible to represent ~1 million trucks by a single object, allowing them to be simulated efficiently.

**Observation:** A piece of intelligence (processed or unprocessed) which the US has collected. It can correspond to a TLO, but it can also represent e.g. a satellite imagery tile that may not contain a TEL or TLO. Observations are marked with what detection modality created the observation (EO satellite, ground sensor, etc.), as well as when the observation occurred.

**File:** A File represents the information the US has collected about a given TEL. The US maintains one File for every TEL, and assigns Observations to a file when it believes them to correspond to it. When it comes time to assess the potential for a nuclear strike, it is the Observations in the Files which are used to do so.

**intelligience.py**
Contains an Intelligence class, representing the efforts of the US intelligence apparatus to find and track TELs. It consists of a pipeline of operations:

> (1) A series of Observer classes (corresponding to different detection modalities) generate raw observations.
>
> (2) Corresponding Analyzer classes process these observations, attempting to determine which correspond to TELs.
>
> (3) A Tracker class takes the observations output from the Analyzers, and attempts to pair them to TEL files.
>
> (4) Finally, an assess() function is called to judge the odds of a successful first strike based on the data in the TEL files.

**observer.py**
Contains several observer classes, all implementing an abstract Observer interface. Each Observer looks at the set of TLOs, and determines which it is able to observe at the current time, based on factors like cloud cover, daylight, TEL state, SAR satellite passes, and so forth. It then adds (a large amount of) observation objects representing raw sensor data which does not correspond to a TLO or TEL.

Contains implementations for EO and SAR satellites, offshore aircraft equipped with SAR sensors, signals intelligence with a low chance to detect TELs not practicing emissions control, and ground sensors with a good likelihood of detecting TELs entering or leaving the base.

**analyzer.py**
Contains implementations of an abstract Analyzer interface. An Analyzer takes a series of observations as input each time step, and then outputs processed observations once it has finished processing them. The operations which are output are only the ones which are believed (by the possibly fallible analysis process) to correspond to real TELs. In reality, some of them may be non-TEL TLOs, such as trucks or decoys.

The ImageryAnalyzer models an AI-assisted image recognition process. It is assumed that large quantities of image data are processed first by an ML algorithm with a given false positive and false negative rate, and then by human analysts. Only a certain quantity of human analysts are available, so how quickly they are able to process a batch of data will depend crucially on how many non-TEL objects can be filtered out by the AI system.

**tracker.py**
Contains implementation of the Tracker interface.

**assessor.py**
Contains the implementation of the assess() method. This is called at each timestep to determine whether a first strike is possible. It works by first collecting the area around each TEL which must be destroyed, which varies based on how precisely the TEL's location is known. This area is calculated several times, based on the flight time delay of different US nuclear assets.

To complete the assessment, missiles in the US nuclear arsenal are matched up against TELs, starting with those whose location is most precisely known (and are thus easiest to destroy). Once enough missiles have been committed to a given TEL to guarantee destruction, it proceeds to the next. Ultimately, either an assignment which destroys all TELs is found, or there is some remaining amount of TELs with the ability to fire back at the US. This number is passed through a formula to determine the probability that at least one warhead gets through US missile defenses.

**renderer.py**
Contains the Renderer class, which is responsible for generating charts and graphs based on the state of the simulation.

# Bibliography


Abbany, Zulfikar. “Modern spy satellites in an age of space wars,” *Deutsche Welle*, August 25, 2020, https://www.dw.com/en/modern-spy-satellites-in-an-age-of-space-wars/a-54691887.

Acemoglu, Daron, and Pascual Restrepo. “Demographics and Automation,” *NBER*, March 2018, https://www.nber.org/papers/w24421.

Acton, James. “The Evolution of Ambiguous Weapons,” *Carnegie Endowment for International Peace*, April 9, 2020, https://carnegieendowment.org/2020/04/09/evolution-of-ambiguous-weapons-pub-81449.

Aftergood, Steven. “EP-3E ARIES,” October 20, 2016, https://fas.org/irp/program/collect/ep-3_aries.htm.

Agrawal, Ajay, John McHale, and Alexander Oett. “Finding Needles in Haystacks: Artificial Intelligence and Recombinant Growth,” in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 1-41. Available online: https://www.nber.org/books/agra-1.

Ahmed, Nur, and Muntasir Wahed. “The De-Democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research,” *arXiv*, October 22, 2020, https://arxiv.org/ftp/arxiv/papers/2010/2010.15581.pdf.

Akkaya, Ilge, Marcin Andrychowicz, Maciek Chociej, Mateusz Litwin, Bob McGrew, Arthur Petron, Alex Paino, Matthias Plappert, Glenn Powell, Raphael Ribas, Jonas Schneider, Nikolas Tezak, Jerry Tworek, Peter Welinder, Lilian Weng, Qiming Yuan, Wojciech Zaremba, and Lei Zhang,.”Solving Rubik’s Cube with a Robot Hand,” *arXiv*, October 16, 2019, https://arxiv.org/pdf/1910.07113.pdf.

Ali, Idrees, and Phil Stewart. “Pentagon concerned by China’s nuclear ambitions, expects warheads to double,” *Reuters*, September 1, 2020, https://www.reuters.com/article/us-usa-china-military-nuclear/pentagon-concerned-by-chinas-nuclear-ambitions-expects-warheads-to-double-idUSKBN25S5MB.

Allen, Greg, and Taniel Chen. “Artificial Intelligence and National Security,” *Harvard Kennedy School*, July 2017, https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf.

Allen, Greg. “Understanding AI Technology,” *DoD Joint Artificial Intelligence Center*, 2020, https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf.

Allen, Greg. “Understanding China’s AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security,” *CNAS*, February 6, 2019, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

Allison, Graham. “Is China Beating America to AI Supremacy?,” *The National Interest*, December 22, 2019, https://nationalinterest.org/feature/china-beating-america-ai-supremacy-106861.

Amodei, Dario, and Danny Hernandez. “AI and Compute,” *OpenAI*, May 16, 2018, https://openai.com/blog/ai-and-compute/.

Amodei, Dario, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mane. “Concrete Problems in AI Safety,” *arXiv*, July 25, 2016, https://arxiv.org/pdf/1606.06565.pdf.

Andersen, Ross. “The Panopticon Is Already Here,” *The Atlantic*, 2020, https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/.

Andrade, Tonio. *The Gunpowder Age: China, Military Innovation, and the Rise of the West in World History* (Princeton: Princeton University Press, 2016), 29.

Andrenelli, Andrea, Julien Gourdon, Yuki Matsumoto, Taku Nemoto, Jehan Sauvage, and Christian Steidl. "Measuring distortions in international markets: The semiconductor value chain," *OECD Trade Policy Papers* 234 (2019), https://www.oecd-ilibrary.org/docserver/8fe4491d-en.pdf?.

Andronov, A. "American signals intelligence satellites in geosynchronous orbit," *Foreign Military Review* 12, trans. Allen Thomson (1993).

Atherton, Kelsey D. "To understand autonomous weapons, think about electronic warfare," *C4ISRNET*, November 15, 2018, https://www.c4isrnet.com/electronic-warfare/2018/11/15/to-understand-autonomous-weapons-think-about-electronic-warfare/.

Axe, David. "7 Secret Ways America's Stealth Armada Stays Off the Radar," *WIRED*, 2012, https://www.wired.com/2012/12/steath-secrets/?pid=1688.

Bae, Ho, Jaehee Jang, Dahuin Jung, Hyemi Jang, Heonseok Ha, and Sungroh Yoon. "Security and Privacy Issues in Deep Learning," *arXiv*, December 6, 2018, https://arxiv.org/pdf/1807.11655.pdf.

Barton, Dominic, Jonathan Woetzel, Jeongmin Seong, and Qinzheng Tian. "Artificial Intelligence: Implications for China," *McKinsey*, April 2017, https://www.mckinsey.com/featured-insights/china/artificial-intelligence-implications-for-china.

Batra, Gaurav, Zach Jacobson, Siddarth Madhav, Andrea Queirolo, and Nick Santhanam. "Artificial-intelligence hardware: New opportunities for semiconductor companies," *McKinsey*, 2019.

Bausch, Andrew W. "Coup-Proofing and Military Inefficiencies: An Experiment," *International Interactions* 44, no. 1 (2018): 1-32.

Berner, Christopher, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, Rafal Józefowicz, Scott Gray, Catherine Olsson, Jakub Pachocki, Michael Petrov, Henrique Pondé de Oliveira Pinto, Jonathan Raiman, Tim Salimans, Jeremy Schlatter, Jonas Schneider, Szymon Sidor, Ilya Sutskever, Jie Tang, Philip Wolski, Susan Zhang. "Dota 2 with Large Scale Deep Reinforcement Learning," *arXiv*, December 13, 2019, https://arxiv.org/pdf/1912.06680.pdf.

Betz, David, and Hugo Stanford-Tuck. "The City Is Neutral: On Urban Warfare in the 21st Century," *TNSR* 2.4 (2019): 60-87.

Biddle, Stephen, and Ivan Oelrich. "Future Warfare in the Western Pacific: Technical Appendix," *Harvard Dataverse*, July 20, 2016, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/GK6PR2, 1-4.

Biddle, Stephen, and Ivan Oelrich. "Future Warfare in the Western Pacific: Chinese Antiaccess/Area Denial, US AirSea Battle, and Command of the Commons in East Asia," *International Security* 41.1 (2016), 46.

Biddle, Stephen. *Military Power: Explaining Victory and Defeat in Modern Battle* (Princeton: Princeton University Press, 2004).

Bin, Li. "Paper Tiger with Whitened Teeth," *China Security* (2006), 78-89. Available online: https://www.issuelab.org/resources/437/437.pdf.

Bin, Li. "Tracking Chinese Strategic Mobile Missiles," *Science and Global Security* 15.1 (2007).

Bitzinger, Richard A. "The Globalization of the Arms Industry: The Next Proliferation Challenge," *International Security* 19.2 (1994), 170-198.

Blank, Steve. “The Chip Wars of the 21st Century,” *War on the Rocks*, 2020, https://warontherocks.com/2020/06/the-chip-wars-of-the-21st-century/.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. “Are Ideas Getting Harder to Find?”, *American Economic Review* 110.4 (2020: 1104-144.

Bohr, Mark. “A 30 Year Retrospective on Dennard's MOSFET Scaling Paper,” 2007, *Solid-State Circuits Society*, http://www.eng.auburn.edu/~agrawvd/COURSE/READING/LOWP/Boh07.pdf.

Bonawitz, Keith, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konecny, Stefano Mazzocchi, H. Brendan McMahan, Timon Van Overveldt, David Petrou, Daniel Ramage, and Jason Roselander. “Towards Federated Learning at Scale: System Design,” *arXiv*, March 22, 2019, https://arxiv.org/pdf/1902.01046.pdf.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies* (Oxford University Press: UK, 2014).

Boulanin, Vincent (ed.). *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume I: Euro-Atlantic Perspectives* (SIPRI, May 2019), https://fas.org/sgp/crs/natsec/R45178.pdf.

Boulanin, Vincent, and Maaike Verbruggen. “Mapping the Development of Autonomy in Weapon Systems,” *SIPRI*, 2017, https://www.sipri.org/publications/2017/other-publications/mappingdevelopment-autonomy-weapon-systems.

Bowers, Ian, and Henrik Stalhane Hiim. “Conventional Counterforce Dilemmas: South Korea's Deterrence Strategy and Stability on the Korean Peninsula,” *International Security* 45.3 (2020/21), 23-4, https://doi.org/10.1162/isec_a_00399.

Bracken, Paul. “The Intersection of Cyber and Nuclear War,” *Strategy Bridge*, January 17, 2017, https://thestrategybridge.org/the-bridge/2017/1/17/the-intersection-of-cyber-and-nuclear-war.

Brahma, Pratik Prabhanjan, Dapeng Wu, and Yiyuan She. “Why Deep Learning Works: A Manifold Disentanglement Perspective,” *IEEE Transactions on Neural Networks and Learning Systems* 27.10 (2016): 1997-2008.

Broersma, Matthew. “Chinese Chip Giant Fujian Jinhua To Cease Operations After US Ban,” *Silicon*, January 2019, https://www.silicon.co.uk/workspace/chinese-chip-giant-fujian-jinhua-to-cease-dram-production-after-us-ban-240903.

Brooks, Stephen G. *Producing Security: Multinational Corporations, Globalization, and the Changing Calculus of Conflict* (Princeton: Princeton University Press, 2006).

Brose, Christian. *The Kill Chain: Defending America in the Future of High-Tech Warfare* (New York: Hachette Books, 2020).

Brown, Noam, Adam Lerer, Sam Gross, and Tuomas Sandholm. “Deep Counterfactual Regret Minimization,” *arXiv*, May 22, 2019, https://arxiv.org/pdf/1811.00164.pdf.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. “Language Models are Few-Shot Learners,” *arXiv*, July 22, 2020, https://arxiv.org/abs/2005.14165, 27.

Brundage, Miles, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, Paul Scharre, Thomas Zeitzoff, Bobby Filar, Hyrum Anderson, Heather Roff, Gregory C. Allen, Jacob Steinhardt,

Carrick Flynn, Sean O hEigeartaigh, Simon Beard, Hadyn Belfield, Sebastian Farquhar, Clare Lyle, Rebecca Crootof, Owain Evans, Michael Page, Joanna Bryson, Roman Yapolsiky, and Dario Amodei. "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," *arXiv*, February 2018, https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf.

Buchanan, Ben, and Taylor Miller, "Machine Learning for Policymakers: What It Is and Why It Matters," *Belfer*, June 2017, https://www.belfercenter.org/sites/default/files/files/publication/MachineLearningforPolicymakers.pdf.

Buchanan, Ben. *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (UK: Oxford University Press, 2016).

Butler, Amy, and Bill Sweetman. "Secret New UAS Shows Stealth, Efficiency Advances," *Aviation Week*, December 6, 2013.

Button, Robert Warren. "Artificial Intelligence and the Military," *RAND*, September 7, 2017, https://www.rand.org/blog/2017/09/artificial-intelligence-and-the-military.html.

Callaway, Ewen. "'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures," *Nature*, November 30, 2020, https://www.nature.com/articles/d41586-020-03348-4.

Cardillo, Robert. "Small Satellites – Big Data," NGA, August 7, 2017, https://www.nga.mil/MediaRoom/Speeches

Caverley, Jonathan D. "United States Hegemony and the New Economics of Defense," *Security Studies*, 16.4 (2007), 598–614.

Chang, Ben Angel. "AI and US-China Relations," in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, ed. Nicholas Wright (DoD SMA: December 2018).

Chang, Chia-Chien, and Alan H. Yang. "Weaponized Interdependence: China's Economic Statecraft and Social Penetration Against Taiwan," *Orbis* 64.2 (2020): 312-33.

Chase, Michael S., Andrew S. Erickson, and Christopher Yeaw. "Chinese Theater and Strategic Missile Force Modernization and its Implications for the United States," *Journal of Strategic Studies* 32, no. 1 (2009): 67-114;

Chen, Lei. "Curse of Dimensionality," in *Encyclopedia of Database Systems*, ed. Ling Liu and M. Tamer Ozsu (Boston: Springer, 2009).

Chen, Xinyun, Chang Liu, Bo Li, Kimberly Lu, and Dawn Song. "Targeted Backdoor Attacks on Deep Learning Systems Using Data Poisoning," *arXiv*, December 15, 2017, https://arxiv.org/pdf/1712.05526.pdf.

Cheng, Victor, "Modern War on an Ancient Battlefield: The Diffusion of American Military Technology and Ideas in the Chinese Civil War, 1946-1949," *Modern China* 35.1 (2009), 38-64.

Chesney, Robert, and Danielle K. Citron. "Disinformation on Steroids," *CFR*, October 16, 2018, https://www.cfr.org/report/deep-fake-disinformation-steroids.

Chessen, Matt. "The MADCOM Future," *The Atlantic Council*, 2017, https://www.atlanticcouncil.org/wp-content/uploads/2017/09/The_MADCOM_Future_RW_0926.pdf.

Chin, Josh, and Liza Lin. "China's All-Seeing Surveillance State Is Reading Its Citizens' Faces," *The Wall Street Journal*, June 26, 2017, https://www.wsj.com/articles/the-all-seeing-surveillance-state-feared-in-the-west-is-areality-in-china-1498493020?

Chin, Josh. “About to Break the Law? Chinese Police Are Already On To You,” *The Wall Street Journal*, April 16, 2019, https://www.wsj.com/articles/china-said-to-deploy-big-data-for-predictive-policing-in-xinjiang-1519719096.

China State Council. “A Next Generation Artificial Intelligence Development Plan,” July 20, 2017, translated by New America, https://www.newamerica.org/documents/1959/translation-fulltext-8.1.17.pdf

Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, Rob Fergus. “Intriguing properties of neural networks,” *arXiv*, February 19, 2014, https://arxiv.org/pdf/1312.6199.pdf.

Clark, Harry, W. Clark McFadden II, and Jeanine McGuinness. “The New 'Uniquely Required' Standard of U.S. Export Licensing Policy for SMIC,” *JD Supra*, 2021, https://www.jdsupra.com/legalnews/the-new-uniquely-required-standard-of-u-2185333/.

Clark, Robert. “Samsung, TSMC $130B plunge underlines the great chip divide,” Light Reading, 2021, https://www.lightreading.com/asia/samsung-tsmc-$130b-plunge-underlines-great-chip-divide/d/d-id/768515.

Clarke, Arthur C. “Superiority,” *The Magazine of Fantasy and Science Fiction* 2.4 (1951), 3-12. Available online: http://nob.cs.ucdavis.edu/classes/ecs153-2019-04/readings/superiority.pdf.

Clary, Christopher, and Vipin Narang. “India’s Counterforce Temptations: Strategic Dilemmas, Doctrine, and Capabilities,” *International Security* 43.3 (2019), 7-52.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern. “The Impact of Artificial Intelligence on Innovation,” forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 1-40. Available online: https://www.nber.org/books/agra-1.

Cohen, Danielle F.S., and Jonathan Kirshner. “The Cult of Energy Insecurity and Great Power Rivalry Across the Pacific,” in *The Nexus of Economics, Security, and International Relations in East Asia*, eds. Avery Goldstein and Edward D. Mansfield (US: Stanford University Press, 2012), 144-76.

Collins, Gabriel B., and William S. Murray. “No Oil for the Lamps of China?”, Naval War College Review 61.2 (2008): 1-17. Available online: https://digital-commons.usnwc.edu/nwc-review/vol61/iss2/10/.

Cote, Owen R. “Assessing the Undersea Balance Between the U.S. and China,” *SSP Working Paper*, 2011, https://www.usni.org/sites/default/files/inline-files/Undersea%20Balance%20WP11-1.pdf.

Cote, Owen R. “The Third Battle: Innovation in the U.S. Navy's Silent Cold War Struggle with Soviet Submarines,” *Naval War College Newport Papers* 16 (2003), https://digital-commons.usnwc.edu/newport-papers/38/.

Cox, Jessica, and Heather Williams. “The Unavoidable Technology: How Artificial Intelligence Can Strengthen Nuclear Stability,” *The Washington Quarterly*, 44.1 (2021): 69-85.

Cramton, Catherine Durnell. “Insights for Culture and Psychology from the Study of Distributed Work Teams,” in *Handbook of Advances in Culture and Psychology, Volume 6*, ed. Michele J. Gelfand, Chi-yue Chiu, and Ying-yi Hong (UK: Oxford University Press, 2015).

Cummings, M. L. “Artificial Intelligence and the Future of Warfare,” *Chatham House*, January 2017, https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf.

Cunningham, Fiona S., and M. Taylor Fravel, “Assuring Assured Retaliation: China’s Nuclear Posture and U.S.-China Strategic Stability,” *International Security* 40.2 (2015): 7-50.

Dafoe, Allan, and Stuart Russell. “Yes, We Are Worried About the Existential Risk of Artificial Intelligence,” *MIT Technology Review*, November 2, 2016, https://www.technologyreview.com/s/602776/yes-we-areworried-about-the-existential-risk-of-artificial-intelligence/.

Dafoe, Allan. “On Technological Determinism: A Typology, Scope Conditions, and a Mechanism,” *Science, Technology, and Human Values* 40.6 (2015), 1047–76.

Danzig, Richard. “An Irresistible Force Meets a Moveable Object: The Technology Tsunami and the Liberal Order,” *Lawfare Research Paper Series* 5.1 (2017), https://assets.documentcloud.org/documents/3982439/Danzig-LRPS1.pdf, 4-7.

Davis, Lynn Etheridge, and Warner R. Schilling. “All You Ever Wanted to Know about MIRV and ICBM Calculations but Were Not Cleared to Ask,” *The Journal of Conflict Resolution* 17.2 (1973).

Davis, Zachary S. “Artificial Intelligence on the Battlefield: An Initial Survey of Potential Implications for Deterrence, Stability, and Strategic Surprise,” *Center for Global Security Research*, March 2019, https://cgsr.llnl.gov/content/assets/docs/CGSR-AI_BattlefieldWEB.pdf.

de Mesquita, Bruce Bueno, Alastair Smith, Randolph M. Siverson, and James D. Morrow. *The Logic of Political Survival* (US: MIT Press, 2003).

Debs, Alexandre, and Nuno P. Monteiro. *Nuclear Politics: The Strategic Causes of Proliferation* (UK: Cambridge University Press, 2017).

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *arXiv*, May 24, 2019, https://arxiv.org/pdf/1810.04805.pdf.

Devlin, Kat, and Christine Huang. “In Taiwan, Views of Mainland China Mostly Negative,” *Pew Research*, May 12, 2020, https://www.pewresearch.org/global/2020/05/12/in-taiwan-views-of-mainland-china-mostly-negative/.

Devries, Phoebe M. R., Fernanda Viegas, Martin Wattenberg, and Brendan J. Meade. “Deep learning of aftershock patterns following large earthquakes,” *Nature* 560 (2018): 632-4. Available online: https://www.nature.com/articles/s41586-018-0438-y.

DiMascio, Jen. “Unmasking the RQ-180,” *Aviation Week & Space Technology*, December 6, 2013.

Dowlin, Nathan, Ran Gilad-Bachrach, Kim Laine, Kristin Lauter, Michael Naehrig, and John Wernsing. “CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy,” in *Proceedings of the 33rd International Conference on Machine Learning* (New York, NY: JMLR, 2016), 1-10. Available online: http://proceedings.mlr.press/v48/gilad-bachrach16.pdf.

Doyle, Michael W. *Liberal Peace: Selected Essays* (New York: Routledge, 2011).

Drexler, K. Eric. “Reframing Superintelligence: Comprehensive AI Services as General Intelligence,” *Future of Humanity Institute*, 2019, https://www.fhi.ox.ac.uk/wp-content/uploads/Reframing_Superintelligence_FHI-TR-2019-1.1-1.pdf.

Drezner, Daniel W. “Technological change and international relations,” *International Relations* 33.2 (2019).

Drezner, Daniel W. “What if AI is just BS?,” *The Washington Post*, May 1, 2019, https://www.washingtonpost.com/outlook/2019/05/01/what-if-ai-is-just-bs/?

Drezner, Daniel W., Henry Farrell, and Abraham L. Newman (eds.). *The Uses and Abuses of Weaponized Interdependence* (Brookings, 2021).

Dutton, Tim, Brent Barron, and Gaga Boskovic. "Building an AI World: Report on National and Regional AI Strategies," *CIFAR*, 2018, https://www.cifar.ca/docs/default-source/ai-society/buildinganaiworld_eng.pdf.

Easton, Ian. *The Chinese Invasion Threat: Taiwan's Defense and American Strategy in Asia* (US: Eastbridge Brooks, 2019).

Epstein, Benjamin, and Roy H. Olsson III. "Physical Signal Classification Via Deep Neural Networks," *arXiv*, November 15, 2018, https://arxiv.org/abs/1811.06349.

Erickson, Andrew S., and David D. Yang. "Using the Land to Control the Sea? Chinese Analysts Consider the Antiship Ballistic Missile," *Naval War College Review* 62.4 (2009), 53-86.

Erickson, Andrew S., Evan Braden Montgomery, Craig Neuman, Stephen Biddle, and Ivan Oelrich. "Correspondence: How Good Are China's Antiaccess/Area-Denial Capabilities?" *International Security* 41.4 (Spring 2017): 202-13.

Esmaeilzadeh, Hadi, Emily Blem, Renée St. Amant, Karthikeyan Sankaralingam, and Doug Burger. "Dark silicon and the end of multicore scaling," paper presented at the *38th Annual International Symposium on Computer Architecture*, San Jose, CA, 2011.

Etzioni, Oren. "No, the Experts Don't Think Superintelligent AI is a Threat to Humanity," *MIT Technology Review*, September 20, 2016, https://www.technologyreview.com/s/602410/no-the-experts-dont-think-superintelligent-ai-is-a-threat-to-humanity.

Farrell, Henry, and Abraham L. Newman. "Weaponized Interdependence: How Global Economic Networks Shape State Coercion," *International Security* 44 (2019).

Feaver, Peter D. *Armed Servants: Agency, Oversight, and Civil-Military Relations* (Cambridge: Harvard University Press, 2003).

Feaver, Peter. "Crisis as Shirking: An Agency Theory Explanation of the Souring of American Civil-Military Relations," *Armed Forces and Society* 24. 3 (1998): 407-34.

Feaver, Peter. "The Civil-Military Problematique: Huntington, Janowitz, and the Question of Civilian Control," *Armed Forces and Society* 23.2 (1996): 149-78.

Fefferman, Charles, Sanjoy Mitter, and Hariharan Narayanan. "Testing the Manifold Hypothesis," *Journal of the American Mathematical Society* 29.4 (2016): 983-1049. Available online: http://www.mit.edu/~mitter/publications/1

Feldstein, Steven. "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression," *Journal of Democracy* 30 (2019).

Fitzpatrick, Mark. "Artificial Intelligence and Nuclear Command and Control," *Survival* 61.3 (2019): 81-92.

Flournoy, Michèle A., Avril Haines, and Gabrielle Chefitz. "Building Trust through Testing: Adapting DOD's Test & Evaluation, Validation & Verification (TEVV) Enterprise for Machine Learning Systems, including Deep Learning Systems," *WestExec Advisors*, 2020, https://cset.georgetown.edu/wp-content/uploads/Building-Trust-Through-Testing.pdf.

Fravel, M. Taylor, and Evan S. Medeiros. "China's Search for Assured Retaliation: The Evolution of Chinese Nuclear Strategy and Force Structure," *International Security* 35.2 (2010), 74-5.

Fravel, M. Taylor. *Active Defense: China's Military Strategy Since 1949* (NJ: Princeton University Press, 2019).

Friedberg, Aaron. *A Contest for Supremacy: China, America, and the Struggle for Mastery in Asia* (New York: W. W. Norton & Company, 2011).

Friedberg, Aaron. *Beyond Air-Sea Battle: The Debate Over US Military Strategy in Asia* (USA: Routledge, 2014).

Friedman, Thomas. *The World is Flat: A Brief History of the Twenty-first Century* (New York: Farrar, Strauss, and Giroux, 2005).

Fujita, Masahisa, and Paul Krugman. "The new economic geography: Past, present and the future," *Papers in Regional Science* 83 (2005): 139-64.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables, *The Spatial Economy: Cities, Regions, and International Trade* (MA: MIT Press, 2001).

Fuller, Douglas B. "Growth, Upgrading, and Limited Catch-Up in China's Semiconductor Industry," in *Policy, Regulation and Innovation in China's Electricity and Telecom Industries*, eds. Loren Brandt and Thomas G. Rawski (Cambridge University Press, 2019).

Fulton, Scott. "How hyperscale data centers are reshaping all of IT," *ZDNet*, April 5, 2019, https://www.zdnet.com/article/how-hyperscale-data-centers-are-reshaping-all-of-it/.

Gai, Silvano. "Dennard Scaling," March 17, 2020, https://silvanogai.github.io/posts/dennard/.

Gamble, Chris, and Jim Gao. "Safety-first AI for autonomous data centre cooling and industrial control," *DeepMind*, August 17, 2018, https://deepmind.com/blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control.

Garfinkel, Ben, and Allan Dafoe. "Artificial Intelligence, Foresight, and the Offense-Defense Balance," *War on the Rocks*, December 19, 2019, https://warontherocks.com/2019/12/artificial-intelligence-foresight-and-the-offense-defense-balance/.

Gartzke, Erik. "Blood and robots: How remotely piloted vehicles and related technologies affect the politics of violence," *Journal of Strategic Studies* (2019). Available online: https://www.tandfonline.com/doi/full/10.1080/01402390.2019.1643329.

Gavin, Francis J. "Strategies of Inhibition: U.S. Grand Strategy, the Nuclear Revolution, and Nonproliferation," *International Security* 44.2 (2019), 185-92.

Gavin, Francis. *Nuclear Statecraft: History and Strategy in America's Atomic Age* (US: Cornell University, 2012).

Geist, Edward, and Andrew J. Lohn. "How Might Artificial Intelligence Affect the Risk of Nuclear War?," *RAND*, 2018, https://www.rand.org/pubs/perspectives/PE296.html.

Gerschenkron, Alexander. *Economic Backwardness in Historical Perspective: A Book of Essays* (MA: Belknap, 1962).

Gholz, Eugene, and Daryl G. Press. "Enduring Resilience: How Oil Markets Handle Disruptions," *Security Studies* 22.1 (2013), 139-47l.

Gholz, Eugene, and Daryl G. Press. "The effects of wars on neutral countries: Why it doesn't pay to preserve the peace," *Security Studies* 10.4 (2001): 1-57.

Gholz, Eugene. "Globalization, Systems Integration, and the Future of Great Power War," *Security Studies* 16.4 (2007): 615-36.

Ghosh, Dipayan, and Ben Scott. "Digital Deceit: The Technologies Behind Precision Propaganda on the Internet," *New America*, January 23, 2018, https://www.newamerica.org/public-interest-technology/policypapers/digitaldeceit/.

Gibbons, Rebecca Davis, and Matthew Kroenig. "Reconceptualizing nuclear risks: Bringing deliberate nuclear use back in," *Comparative Strategy* 35, no. 5 (2016): 407-22.

Gilli, Andrea, and Mauro Gilli. "The Diffusion of Drone Warfare? Industrial, Organizational, and Infrastructural Constraints," *Security Studies* 25.1 (2016): 50-84.

Gilli, Andrea, and Mauro Gilli. "Why China Has Not Caught Up Yet: Military-Technological Superiority and the Limits of Imitation, Reverse Engineering, and Cyber Espionage," *International Security* 43.3 (2018/19): 141-89.

Gilpin, Leilani H., David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. "Explaining Explanations: An Overview of Interpretability of Machine Learning," *arXiv*, February 3, 2019, https://arxiv.org/pdf/1806.00069.pdf.

Gilpin, Robert. *War and Change in World Politics* (Cambridge: Cambridge University Press, 1981).

Glaser, Charles L., and Steve Fetter. "Should the United States Reject MAD? Damage Limitation and U.S. Nuclear Strategy toward China," *International Security* 41.1 (2016): 49-98.

Goddard, Kate, Abdul Roudsari, and Jeremy C. Wyatt. "Automation bias: a systematic review of frequency, effect mediators, and mitigators," *Journal of the American Medical Informatics Association* 19 (2012): 121-7.

Goldfarb, Avi, and Daniel Trefler, "AI and International Trade," forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 8-9. Available online: https://www.nber.org/books/agra-1.

Goldfein, Jocelyn, and Ivy Nguyen. "Data is not the new oil," *TechCrunch*, March 27, 2018, https://techcrunch.com/2018/03/27/data-is-not-the-new-oil/.

Goldman, Emily O. "Cultural foundations of military diffusion," *Review of International Studies* 32.1: 69-91.

Goldman, Emily O., and Leslie C. Eliason (eds.). *The Diffusion of Military Technology and Ideas* (California: Stanford University Press, 2003).

Goldman, Emily O., and Richard B. Andres. "Systemic effects of military innovation and diffusion*," Security Studies* 8 (1999).

Goldstein, Avery. "First Things First: The Pressing Danger of Crisis Instability in U.S.-China Relations," *International Security* 37.4 (2013): 49-89.

Goldstein, Seth. "Hybrid Forecasting Competition (HFC)," *IARPA*, accessed May 10, 2019, https://www.iarpa.gov/index.php/research-programs/hfc.

Gompert, David C., Astrid Stuth Cevallos, and Cristina L. Garafola. "War with China: Thinking Through the Unthinkable," *RAND*, 2016, https://www.rand.org/pubs/research_reports/RR1140.html.

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative Adversarial Nets," *arXiv*, June 10, 2014, https://arxiv.org/pdf/1406.2661.pdf.

Goodfellow, Ian, Nicolas Papernot, Sandy Huang, Rocky Duan, PIeter Abbeel, and Jack Clark. "Attacking Machine Learning with Adversarial Examples," *OpenAI*, February 24, 2017.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning* (MA: MIT Press, 2016).

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. "When Will AI Exceed Human Performance? Evidence from AI Experts," *Journal of Artificial Intelligence Research* 62 (2018), 729-54.

Graham, William. "Russia's Soyuz-2-1b launches missile detection satellite," *NASASpaceFlight.com*, May 22, 2020, https://www.nasaspaceflight.com/2020/05/russias-soyuz-2-1b-missile-detection-satellite/.

Grant, Rebecca. "The Perils of Chrome Dome," *Air Force Magazine*, 2011, https://www.airforcemag.com/article/0811dome/.

Green, Brendan Rittenhouse, and Austin Long. "Conceal or Reveal? Managing Clandestine Military Capabilities in Peacetime Competition," *International Security* 44.3 (2019/2020), 48-83.

Green, Brendan Rittenhouse, Austin Long, Matthew Kroenig, Charles L. Glaser, and Steve Fetter. "Correspondence: The Limits of Damage Limitation," *International Security* 42.1 (2017): 193-207.

Green, Brendan Rittenhouse. *The Revolution That Failed: Nuclear Competition, Arms Control, and the Cold War* (UK: Cambridge University Press, 2020).

Grigorescu, Sorin, Bogdan Trasnea, Tiberiu Cocias, and Gigel Macesanu, "A Survey of Deep Learning Techniques for Autonomous Driving," *arXiv*, October 17, 2019, https://arxiv.org/pdf/1910.07738.pdf.

Haas, Mark L. *The Ideological Origins of Great Power Politics, 1789-1989* (Ithaca: Cornell University Press, 2005).

Hallex, Matthew A., and Travis S. Cottom. "Proliferated Commercial Satellite Constellations: Implications for National Security," *Joint Forces Quarterly* 97 (2020), 24. Available online: https://ndupress.ndu.edu/Portals/68/Doc

Halterman, Andy, and Rachel Tecott. "The Case for Campaign Analysis: A Method for Studying Military Operations," *International Security* 45 (2021): 44–83.

Hammes, T. X.,. "The Future of Warfare: Small, Many, Smart vs. Few & Exquisite?", *War on the Rocks*, July 16, 2014, https://warontherocks.com/2014/07/the-future-of-warfare-small-many-smart-vs-few-exquisite/.

Hanacek, Joseph. "The Perfect Can Wait: Good Solutions to the 'Drone Swarm' Problem," August 14, 2018, *War on the Rocks*, https://warontherocks.com/2018/08/the-perfect-can-wait-good-solutions-to-the-drone-swarm-problem/.

Hannas, William C., and Huey-meei Chang. "China's Access to Foreign AI Technology: An Assessment," *CSET*, 2019, https://cset.georgetown.edu/research/chinas-access-to-foreign-ai-technology/.

Harer, Jacob A., Louis Y. Kim, Rebecca L. Russell, Onur Ozdemir, Leonard R. Kosta, Akshay Rangamani, Lei H. Hamilton, Gabriel I. Centeno, Jonathan R. Key, Paul M. Ellingwood, Erik Antelman, Alan Mackay, Marc W. McConley, Jeffrey M. Opper, Peter Chin, and Tomo Lazovich. "Automated software vulnerability detection with machine learning," *arXiv*, August 2, 2018, https://arxiv.org/pdf/1803.04497.pdf.

Harper, Jon. "Artificial Intelligence to Sort Through ISR Data Glut," *National Defense*, January 16, 2018, https://www.nationaldefensemagazine.org/articles/2018/1/16/artificial-intelligence-to--sort-through-isr-data-glut.

Haugeland, John. *Artificial Intelligence: The Very Idea* (US: MIT Press, 1985).

Hawkins, Amy. "Beijing's Big Brother Tech Needs African Faces," *Foreign Policy*, July 24, 2018, https://foreignpolicy.com/2018/07/24/beijings-big-brother-tech-needs-african-faces/.

Hawley, John K. “Patriot Wars: Automation and the Patriot Air and Missile Defense System,” *CNAS*, January 2017, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Report-EthicalAutonomy5-PatriotWars-FINAL.pdf?

Heginbotham, Eric, Michael Nixon, Forrest E. Morgan, Jacob L. Heim, Jeff Hagen, Sheng Tao Li, Jeffrey Engstrom, Martin C. Libicki, Paul DeLuca, David A. Shlapak, David R. Frelinger, Burgess Laird, Kyle Brady, and Lyle J. Morris. “The U.S.-China Military Scorecard: Forces, Geography, and the Evolving Balance of Power, 1996–2017,” *RAND*, 2015, https://www.rand.org/pubs/research_reports/RR392.html.

Heginbotham, Eric, Michael S. Chase, Jacob L. Heim, Bonny Lin, Mark R. Cozad, Lyle J. Morris, Christopher P. Twomey, Forrest E. Morgan, Michael Nixon, Cristina L. Garafola, and Samuel K. Berkowitz. “China’s Evolving Nuclear Deterrent: Major Drivers and Issues for the United States,” *RAND*, 2017, https://www.rand.org/pubs/research_reports/RR1628.html.

Hestness, Joel, Sharan Narang, Newsha Ardalani, Gregory Diamos, Heewoo Jun, Hassan Kianinejad, Md. Mostofa Ali Patwary, Yang Yang, and Yanqi Zhou. “Deep Learning Scaling is Predictable, Empirically,” December 1, 2017, *arXiv*, https://arxiv.org/pdf/1712.00409.pdf.

Hewish, Mark. “Reformatting fighter tactics,” *Jane's International Defense Review*, 2007. Available online: https://web.archive.org/web/20070815215345/http://www.textrondefense.com/pdfs/news/jidr06_01.pdf.

Hille, Kathrin. “Trade war forces Chinese chipmaker Fujian Jinhua to halt output,” *Financial Times*, January 28, 2019, https://www.ft.com/content/87b5580c-22bf-11e9-8ce6-5db4543da632.

Hipple, Matthew. “Bring on the Countermeasure Drones,” *USNI*, February 2014, https://www.usni.org/magazines/proceedings/2014/february/bring-countermeasure-drones.

Hoadley, Daniel S., and Kelley M. Sayler. “Artificial Intelligence and National Security,” *Congressional Research Service*, January 30, 2019, https://fas.org/sgp/crs/natsec/R45178.pdf.

Hoffman, D. E. *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy* (Anchor Books: New York, 2009).

Hofstadter, Douglas R. *Gödel, Escher, Bach: An Eternal Golden Braid* (US, Basic Books, 1979).

Holmes, James R. “Taiwan Needs a Maoist Military,” *Foreign Policy*, October 17, 2019, https://foreignpolicy.com/20

Horowitz, Michael C. “Artificial Intelligence, International Competition, and the Balance of Power,” *Texas National Security Review* 1 (2018), https://doi.org/10.15781/T2639KP49.

Horowitz, Michael C. “When speed kills: Lethal autonomous weapon systems, deterrence and stability,” *Journal of Strategic Studies* 42 (2019): 764-88.

Horowitz, Michael C. *The Diffusion of Military Power: Causes and Consequences for International Politics* (New Jersey: Princeton University Press, 2010).

Horowitz, Michael C., Gregory C. Allen, Elsa B. Kania, and Paul Scharre. “Strategic Competition in an Era of Artificial Intelligence,” *CNAS*, 2018, https://www.cnas.org/publications/reports/strategic-competition-in-an-era-of-artificial-intelligence.

Horowitz, Michael C., Joshua A. Schwartz, and Matthew Fuhrmann. “China Has Made Drone Warfare Global,” *Foreign Affairs*, November 20, 2020, www.foreignaffairs.com/articles/china/2020-11-20/china-has-madedrone-warfare-global.

Horowitz, Michael C., Paul Scharre, and Alexander Velez-Green. “A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence,” *arXiv*, 2019, https://arxiv.org/abs/1912.05291.

Horowitz, Michael C., Sarah E. Kreps, and Matthew Fuhrmann. “Separating Fact from Fiction in the Debate over Drone Proliferation,” *International Security* 41.2 (2016), 7-42.

Horowitz, Michael. “Do Emerging Military Technologies Matter for International Politics?”, *Annual Review of Political Science* 23 (2020).

Horowitz, Michael. “Information-Age Weaponry and the Future Shape of Security in East Asia,” *GlobalAsia*, 2011, https://globalasia.org/v6no2/cover/information-age-weaponry-and-the-future-shape-of-security-in-east-asia_michael -horowitz.

Hsu, Jeremy. “Wanted: AI That Can Spy,” *IEEE Spectrum*, November 2017, https://spectrum.ieee.org/aerospace/satellites/wanted-ai-that-can-spy.

Hughes, Llewelyn, and Austin Long. “Is There an Oil Weapon? Security Implications of Changes in the Structure of the International Oil Market,” *International Security* 39.3 (2014/15): 152-89.

Hughes, Llewelyn, and Phillip Y. Lipscy. “The Politics of Energy,” *Annual Review of Political Science* 16.1 (2013), 449–69.

Hunt, Will, Saif Khan, and Dahlia Peterson. “China’s Progress in Semiconductor Manufacturing Equipment: Accelerants and Policy Implications,” *CSET*, March 2021, 10-11.

Hutson, Matthew. “AI researchers allege that machine learning is alchemy,” *Science*, May 3, 2018, https://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy.

Hwang, Tim. “Computational Power and the Social Impact of Artificial Intelligence,” *arXiv*, March 2018, https://arxiv.org/abs/1803.08971.

Hwang, Tim. “Shaping the Terrain of AI Competition,” *CSET*, June 2020, https://cset.georgetown.edu/research/shaping-the-terrain-of-ai-competition/.

Ignatius, David. “The wizards of Armageddon may be back,” *The Washington Post*, 2021, https://www.washingtonpost.com/opinions/2021/05/06/wizards-armageddon-may-be-back/.

Ikenberry, G. John. *A World Safe for Democracy: Liberal Internationalism and the Crises of Global Order* (US: Yale University Press, 2020).

Ilachinski, Andrew. “AI, Robots, and Swarms: Issues, Questions, and Recommended Studies,” *CNA*, January 2017, https://www.cna.org/cna_files/pdf/DRM-2017-U-014796-Final.pdf.

Imbrie, Andrew, Elsa B. Kania, and Lorand Laskai. “The Question of Comparative Advantage in Artificial Intelligence: Enduring Strengths and Emerging Challenges for the United States,” *CSET*, January 20, http://cset.georgetown.edu/wp-content/uploads/CSET-The-Question-of-Comparative-Advantage-in-Artificial-Intelli

Imbrie, Andrew, Ryan Fedasiuk, Catherine Aiken, Tarun Chhabra, and Husanjot Chahal. “Agile Alliances: How the United States and Its Allies Can Deliver a Democratic Way of AI,” *CSET*, February 2020, https://cset.georgetown.edu/research/agile-alliances/.

Jalali, Anahid N., Alexander Schindler and Bernhard Haslhofer. “Understandable Deep Neural Networks for Predictive Maintenance in the Manufacturing Industry,” *ERCIM*, January 22, 2019, https://ercim-news.ercim.eu/en116/r-i/understandable-deep-neural-networks-for-predictive-maintenance-in-the-manufacturing-industry.

Jensen, Benjamin M., Christopher Whyte, and Scott Cuomo. “Algorithms at War: The Promise, Peril, and Limits of Artificial Intelligence,” *International Studies Review* 0 (2019), 4.

Jensen, Benjamin, Scott Cuomo, and Chris Whyte, “Wargaming with Athena: How to Make Militaries Smarter, Faster, and More Efficient with Artificial Intelligence,” *War on the Rocks*, June 5, 2018, https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries-smarter-faster-and-more-efficient-with-artificial-intelligence/.

Jensen, W. G. “The Importance of Energy in the First and Second World Wars,” *The Historical Journal* 11 (1968): 538-54.

Jervis, Robert. *The Illogic of American Nuclear Strategy* (New York: Cornell University Press, 1984).

Jervis, Robert. *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (New York: Cornell University Press, 1989).

Johnson, James. “Artificial Intelligence in Nuclear Warfare: A Perfect Storm of Instability?,” *Washington Quarterly* 43.2 (2020): 197–211.

Johnson, James. “Artificial Intelligence, Drone Swarming and Escalation Risks in Future Warfare,” *The RUSI Journal* 165.2 (2020): 26-36.

Johnson, James. “The end of military-techno Pax Americana? Washington’s strategic responses to Chinese AI-enabled military technology,” *The Pacific Review* (2019), https://www.tandfonline.com/action/showCitFormats?doi=

Johnson, Joseph. “MAD in an AI Future?”, *Lawrence Livermore National Laboratory*, June 14, 2019, https://www.osti.gov/servlets/purl/1527284.

Jones, Benjamin F. “The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?” *Review of Economic Studies* 76.1 (2009), 283–317.

Kahneman, Daniel. *Thinking, Fast and Slow* (US: Farrar, Straus and Giroux, 2011).

Kallenborn, Zachary, and Philipp C. Bleek. “Drones of Mass Destruction: Drone Swarms and the Future of Nuclear, Chemical, and Biological Weapons,” *War on the Rocks*, February 14, 2019, https://warontherocks.com/2019/02/drones-of-mass-destruction-drone-swarms-and-the-future-of-nuclear-chemical-and-biological-weapons/.

Kallenborn, Zachary. “AI Risks to Nuclear Deterrence Are Real,” *War on the Rocks*, October 10, 2019, https://warontherocks.com/2019/10/ai-risks-to-nuclear-deterrence-are-real/.

Kallenborn, Zachary. “Are Drone Swarms Weapons of Mass Destruction?” Center for Strategic Deterrence Studies, 2020, https://media.defense.gov/2020/Jun/29/2002331131/-1/-1/0/60DRONESWARMS-MONOGRAPH.PDF.

Kallenborn, Zachary. “Meet the future weapon of mass destruction, the drone swarm,” *Bulletin of the Atomic Scientists*, April 5, 2021, https://thebulletin.org/2021/04/meet-the-future-weapon-of-mass-destruction-the-drone-swarm/.

Kanaan, Michael. *T Minus AI: Humanity’s Countdown to Artificial Intelligence and the New Pursuit of Global Power* (Texas: BenBella Books, Inc., 2020).

Kania, Elsa B. “‘AI Weapons’ in China’s Military Innovation,” *Brookings*, April 2020, https://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf.

Kania, Elsa B. “Battlefield Singularity: Artificial Intelligence, Military Revolution, and China’s Future Military Power,” *CNAS*, November 28, 2017, https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power.

Kastner, Scott L. “Is the Taiwan Strait Still a Flash Point? Rethinking the Prospects for Armed Conflict between China and Taiwan,” *International Security* 40.3 (2015/16).

Kates-Harbeck, Julian, Alexey Svyatkovskiy, and William Tang, “Predicting disruptive instabilities in controlled fusion plasmas through deep learning,” *Nature* 568 (2019): 526-42.

Keith, Adam. “Is hyperspectral the next Earth observation frontier?,” *SpaceNews*, March 30, 2019, https://spacenews.com/op-ed-is-hyperspectral-the-next-earth-observation-frontier/.

Kelanic, Rosemary A. “Black Gold and Blackmail: The Politics of International Oil Coercion” (PhD diss., University of Chicago, 2012).

Kelly, Kevin. “The Three Breakthroughs That Have Finally Unleashed AI on the World,” *Wired*, October 27, 2014, https://www.wired.com/2014/10/future-of-artificial-intelligence/.

Kemp, R. Scott. “The Nonproliferation Emperor Has No Clothes: The Gas Centrifuge, Supply-Side Controls, and the Future of Nuclear Proliferation,” *International Security* 38 (2014): 39-78, especially 41-4.

Kennedy, Paul. *The Rise and Fall of the Great Powers* (New York: Random House, 1987).

Kent, Glenn A., and David E. Thaler. “First-Strike Stability: A Methodology for Evaluating Strategic Forces,” *RAND*, 1989, https://www.rand.org/pubs/reports/R3765.html.

Khan, Saif M. “Maintaining the AI Chip Competitive Advantage of the United States and its Allies,” *CSET*, 2019, https://cset.georgetown.edu/wp-content/uploads/CSET-Maintaining-the-AI-Chip-Competitive-Advantage-of-the-United-States-and-its-Allies-20191206.pdf.

Khan, Saif M., Alexander Mann, and Dahlia Peterson. “The Semiconductor Supply Chain: Assessing National Competitiveness,” *CSET*, 2021, https://cset.georgetown.edu/wp-content/uploads/The-Semiconductor-Supply-Chain-Issue-Brief.pdf.

Khan, Saif M., and Alexander Mann. “AI Chips: What They Are and Why They Matter,” *CSET*, April 2020, https://cset.georgetown.edu/wp-content/uploads/AI-Chips%E2%80%94What-They-Are-and-Why-They-Matter.pdf.

King, Gary, Jennifer Pan, and Margaret E. Roberts. “How Censorship in China Allows Government Criticism but Silences Collective Expression,” *American Political Science Review* 107 (2013), 326-43.

King, Gary. “Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data,” *International Organization* 55 (2001): 497-507.

Kirshner, Jonathan. “The Changing Calculus of Conflict?”, *Security Studies* 16.4 (2007): 583-597.

Kirshner, Jonathan. “The tragedy of offensive realism: Classical realism and the rise of China,” *European Journal of International Relations* 18.1 (2012): 53-75.

Kissinger, Henry A. “How the Enlightenment Ends,” *The Atlantic*, June 2018, https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

Kissinger, Henry A., Eric Schmidt, and Daniel Huttenlocher. "The Metamorphosis," *The Atlantic*, August 2019, https://www.theatlantic.com/magazine/archive/2019/08/henry-kissinger-the-metamorphosis-ai/592771/.

Kissinger, Henry. *World Order* (USA: Penguin Books Limited, 2014).

Kliman, Daniel M. *Fateful Transitions: How Democracies Manage Rising Powers, from the Eve of World War I to China's Ascendance* (Philadelphia: University of Pennsylvania Press, 2015).

Klinger, Joel, Juan Mateos-Garcia, and Konstantinos Stathoulopoulos. "A narrowing of AI research?," *arXiv*, November 18, 2020, https://arxiv.org/pdf/2009.10385.pdf.

Knack, Stephen, and Philip Keefer. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics* 112 (1997): 1251-88.

Kolyuka, Yu F., N. M. Ivanov, T. I. Afanasieva, and T. A. Gridchina. "Examination of the Lifetime, Evolution, and Re-Entry Features for the 'Molniya' Type Orbits" (paper presented at the 21st International Symposium of Space Flight Dynamics, Toulouse, France, 2009). Available online: https://issfd.org/ISSFD_2009/CollisionRiskII/Kolyuka.pdf.

Konaev, Margarita, Husanjot Chahal, Ryan Fedasiuk, Tina Huang, and Ilya Rahkovsky. "U.S. Military Investments in Autonomy and AI," *CSET*, October 2020, https://cset.georgetown.edu/wp-content/uploads/U.S.-Military-Investments-in-Autonomy-and-AI_Strategic-Assessment-1.pdf.

Konaev, Margarita, Tina Huang, and Husanjot Chahal. "Trusted Partners: Human-Machine Teaming and the Future of Military AI," *CSET*, https://cset.georgetown.edu/research/trusted-partners/.

Konaev, Margarita. "With AI, We'll See Faster Fights, But Longer Wars," *War on the Rocks*, October 29, 2019, https://warontherocks.com/2019/10/with-ai-well-see-faster-fights-but-longer-wars/.

Kovite, Molly. "I, Black Box: Explainable Artificial Intelligence and the Limits of Human Deliberative Processes," *War on the Rocks*, July 5, 2019, https://warontherocks.com/2019/07/i-black-box-explainable-artificial-intelligence-and-the-limits-of-human-deliberative-processes/.

Krafft, P. M., Meg Young, Michael Katell, Karen Huang, and Ghislain Bugingo. "Defining AI in Policy versus Practice," *arXiv*, December 23, 2019, https://arxiv.org/pdf/1912.11095v1.pdf.

Kreps, Sarah, and Miles McCain. "Not Your Father's Bots: AI is Making Fake News Look Real," *Foreign Affairs*, August 2, 2019, www.foreignaffairs.com/articles/2019-08-02/not-your-fathers-bots.

Kristensen, Hans M., and Matt Korda. "Chinese nuclear forces, 2019," *Bulletin of the Atomic Scientists* 75.4 (2019): 171-8.

Kristensen, Hans M., and Matt Korda. "United States nuclear forces, 2020," *Bulletin of the Atomic Scientists* 76.1 (2020): 46-60.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." *Proceedings of the 25th International Conference on Neural Information Processing Systems*, December 2012, 1097–105. Available online: https://dl.acm.org/doi/10.5555/2999134.2999257.

Kroenig, Matthew. *Exporting the Bomb: Technology Transfer and the Spread of Nuclear Weapons* (Ithaca: Cornell University Press, 2010).

Krugman, Paul. "The New Economic Geography, Now Middle-aged," *Regional Studies* 45 (2010): 1-7;

Krugman, Paul. *Geography and Trade* (MA: MIT Press, 1991)

Kuo, Lily. “Taiwan’s citizens battle pro-China fake news campaigns as election nears,” *The Guardian*, December 30, 2019.

Kurakin, Alexey, Ian Goodfellow, Samy Bengio, Yinpeng Dong, Fangzhou Liao, Ming Liang, Tianyu Pang, Jun Zhu, Xiaolin Hu, Cihang Xie, Jianyu Wang, Zhishuai Zhang, Zhou Ren, Alan Yuille, Sangxia Huang, Yao Zhao, Yuzhe Zhao, Zhonglin Han, Junjiajia Long, Yerkebulan Berdibekov, Takuya Akiba, Seiya Tokui, and Motoki Abe. “Adversarial Attacks and Defences Competition,” *arXiv*, March 31, 2018, https://arxiv.org/pdf/1804.00097.pdf.

Kurakin, Alexey, Ian J. Goodfellow, and Samy Bengio. “Adversarial Examples in the Physical World,” *arXiv*, February 11, 2017, https://arxiv.org/pdf/1607.02533.pdf.

Lake, David A. “Open economy politics: A critical review,” *The Review of International Organizations* 4 (2009).

Langbroek, Marco. “A NEMESIS in the sky,” *The Space Review*, October 31, 2016, https://www.thespacereview.com/article/3095/1.

Lanier, Jaron, and Glen Weyl. “AI is an Ideology, Not a Technology,” *WIRED*, March 15, 2020, https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/.

Laskai, Lorand. “Civil-Military Fusion: The Missing Link Between China’s Technological and Military Rise,” *Council on Foreign Relations*, January 29, 2018, https://www.cfr.org/blog/civil-military-fusion-missing-link-between-chinas-technological-and-military-rise.

Lau, Tim. “Predictive Policing Explained,” *Brennan Center for Justice*, April 1, 2020, https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. “Deep learning,” *Nature* 521 (2015): 436-44.

LeCun, Yann. “My take on Ali Rahimi’s ‘Test of Time’ award talk at NIPS,” December 6, 2017, https://www2.isye.gatech.edu/~tzhao80/Yann_Response.pdf.

Lee, John, and Jan-Peter Kleinhans. “Would China Invade Taiwan for TSMC?”, *The Diplomat*, December 15, 2020, https://thediplomat.com/2020/12/would-china-invade-taiwan-for-tsmc/.

Lee, Kai-Fu. *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt, 2018).

Lehman, Joel, Jeff Clune, Dusan Misevic, Christoph Adami, Lee Altenberg, Julie Beaulieu, Peter J. Bentley, Samuel Bernard, Guillaume Beslon, David M Bryson, Patryk Chrabaszcz, Nick Cheney, Antoine Cully, Stephane Doncieux, Fred C. Dyer, Kai Olav Ellefsen, Robert Feldt, Stephan Fischer, Stephanie Forrest, Antoine Frenoy, Christian Gagne, Leni Le Goff, Laura M Grabowski, Babak Hodjat, Frank Hutter, Laurent Keller, Carole Knibbe, Peter Krcah, Richard E. Lenski, Hod Lipson, Robert MacCurdy, Carlos Maestre, Risto Miikkulainen, Sara Mitri, David E. Moriarty, Jean-Baptiste Mouret, Anh Nguyen, Charles Ofria, Marc Parizeau, David Parsons, Robert T. Pennock, William F. Punch, Thomas S. Ray, Marc Schoenauer, Eric Schulte, Karl Sims, Kenneth O Stanley, Francois Taddei, Danesh Tarapore, Simon Thibault, Westley Weimer, Richard Watson, and Jason Yosinski. “The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities,” *arXiv*, August 14, 2018, https://arxiv.org/pdf/1803.03453.pdf.

Levi, Michael. “The Enduring Vulnerabilities of Oil Markets,” *Security Studies* 22.1 (2013), 132–138.

Lewis, Jeffrey, David Joel La Boon, and Decker Eveleth. “China’s Growing Missile Arsenal and the Risk of a ‘Taiwan Missile Crisis,’” *Nuclear Threat Initiative*, November 18, 2020, https://www.nti.org/analysis/articles/chinas-growing-missile-arsenal-and-the-risk-of-a-taiwan-missile-crisis/.

Li, Yuxi. “Deep Reinforcement Learning: An Overview,” *arXiv*, November 26, 2018, https://arxiv.org/abs/1701.07274.

Lichter, Andreas, Max Loffler, and Sebastian Siegloch. “The long-term costs of government surveillance: Insights from Stasi spying in East Germany,” *SOEPpapers on Multidisciplinary Panel Data Research* 865 (2016): 1-60. Available online: https://www.econstor.eu/bitstream/10419/146890/1/869045423.pdf.

Lieber, Keir A., and Daryl G. Press, “The End of MAD? The Nuclear Dimension of U.S. Primacy,” *International Security* 30.4 (2006): 7-44.

Lieber, Keir A., and Daryl G. Press, “The New Era of Counterforce: Technological Change and the Future Deterrence,” *International Security* 41.4 (2017): 9-49. Available online: https://www.mitpressjournals.org/doi/pdf/10.1162/ISEC_a_00273.

Lieber, Keir, and Daryl Press. *The Myth of the Nuclear Revolution: Power Politics in the Atomic Age* (New York: Cornell University, 2020).

Lighthill, James. “Artificial Intelligence: A General Survey,” *Science Research Council*, 1973. Available online: http://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/p001.htm.

Lin, Henry W., Max Tegmark, and David Rolnick. “Why does deep and cheap learning work so well?,” *arXiv*, 2021, https://arxiv.org/abs/1608.08225.

Lind, Jennifer, and Daryl G. Press. “Markets or Mercantilism? How China Secures Its Energy Supplies,” *International Security* 42.4 (2018).

Lin-Greenberg, Erik, Reid Pauly, and Jacquelyn Schneider. “Wargaming for Political Science Research,” *SSRN*, 2021, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3676665.

Lin-Greenberg, Erik. “Remote Controlled Restraint: The Effect of Remote Warfighting Technology on Crisis Escalation” (PhD diss., Columbia University, 2019).

Lin-Greenberg, Erik. “Wargame of Drones: Remotely Piloted Aircraft and Crisis Escalation” (unpublished manuscript), 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3288988.

Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sanchez. “A Survey on Deep Learning in Medical Image Analysis,” *arXiv*, June 4, 2017, https://arxiv.org/pdf/1702.05747.pdf.

Littell, Joe. “Don’t Believe Your Eyes (Or Ears): The Weaponization of Artificial Intelligence, Machine Learning, and Deepfakes,” *War on the Rocks*, October 7, 2019, https://warontherocks.com/2019/10/dont-believe-your-eyes-or-ears-the-weaponization-of-artificial-intelligence-machine-learning-and-deepfakes/.

Lohn, Andrew. “Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity,” *CSET*, 2020, https://cset.georgetown.edu/research/hacking-ai/.

Long, Austin, and Brendan Rittenhouse Green. “Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy,” *Journal of Strategic Studies* 38 (2015): 38-73.

Loss, Rafael, and Joseph Johnson. “Will Artificial Intelligence Imperil Nuclear Deterrence?”, *War on the Rocks*, September 19, 2019, https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/.

Loss, Rafael. “Artificial Intelligence, the Final Piece to the Counterforce Puzzle?”, *Lawrence Livermore National Laboratory*, September 30, 2019, https://www.osti.gov/servlets/purl/1568008.

Lostumbo, Michael J., David R. Frelinger, James Williams, and Barry Wilson. “Air Defense Options for Taiwan: An Assessment of Relative Costs and Operational Benefits,” *RAND*, 2016, https://www.rand.org/pubs/research_reports/RR1051.html, 125-6, 70-1.

Lowther, Adam, and Curtis McGiffin, “America Needs a ‘Dead Hand,’” *War on the Rocks*, August 16, 2019, https://warontherocks.com/2019/08/america-needs-a-dead-hand/.

Luong, Nguyen Cong, Dinh Thai Hoang, Shimin Gong, Dusit Niyato, Ping Wang, Ying-Chang Liang, and Dong In Kim. “Applications of Deep Reinforcement Learning in Communications and Networking: A Survey,” *IEEE Communications Surveys & Tutorials* 21.4 (2019): 3133-74.

Ma, Joy Dantong. “China’s AI Talent Base Is Growing, and then Leaving,” *MacroPolo*, July 30, 2019, https://macropolo.org/chinas-ai-talent-base-is-growing-and-then-leaving/.

Ma, Joy Dantong. “The AI Race Is Wide Open, If America Remains Open,” *Macro Polo*, April 15, 2019, https://macropolo.org/us-china-ai-race-talent/.

Macdonald, Julia, and Jacquelyn Schneider. “Battlefield Responses to New Technologies: Views from the Ground on Unmanned Aircraft,” *Security Studies* 28.2 (2019), 216-49.

MacDonald, Norine, and George Howell. “Killing Me Softly: Competition in Artificial Intelligence and Unmanned Aerial Vehicles,” *PRISM* 8.3 (2019), 102-27.

Marcum, Richard A., Curt H. Davis, Grant J. Scott, and Tyler W. Nivin. “Rapid broad area search and detection of Chinese surface-to-air missile sites using deep convolutional neural networks,” *Journal of Applied Remote Sensing* 11.4 (2017).

Martinez, Antonio Garcia. “No, Data is Not the New Oil,” *WIRED*, February 26, 2019, https://www.wired.com/story/no-data-is-not-the-new-oil/.

Martinic, Gary. “Glimpses of Future Battlefield Medicine – the Proliferation of Robotic Surgeons and Unmanned Vehicles and Technologies,” *Journal of Military and Veterans’ Health* 22 (2014): 4-12.

McCandlish, Sam, Jared Kaplan, and Dario Amodei. “How AI Training Scales,” *OpenAI*, December 14, 2018, https://openai.com/blog/science-of-ai/.

McKinney, Scott Mayer, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, Greg C. Corrado, Ara Darzi, Mozziyar Etemadi, Florencia Garcia-Vicente, Fiona J. Gilbert, Mark Halling-Brown, Demis Hassabis, Sunny Jansen, Alan Karthikesalingam, Christopher J. Kelly, Dominic King, Joseph R. Ledsam, David Melnick, Hormuz Mostofi, Lily Peng, Joshua Jay Reicher, Bernardino Romera-Paredes, Richard Sidebottom, Mustafa Suleyman, Daniel Tse, Kenneth C. Young, Jeffrey De Fauw, and Shravya Shetty. “International evaluation of an AI system for breast cancer screening,” *Nature* 577 (2020), 89-94.

McKinzie, Matthew G., Thomas B. Cochran, Robert S. Norris, and William M. Arkin. “The U.S. Nuclear War Plan: A Time for Change,” *Natural Resources Defense Council*, 2001, https://www.nrdc.org/sites/default/files/us-nuclear-war-plan-report.pdf, 54.

McMahan, H. Brendan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas. “Communication-efficient learning of deep networks from decentralized data,” *arXiv*, February 28, 2017, https://arxiv.org/pdf/1602.05629.pdf.

Mearsheimer, John J., and Stephen M. Walt. “Leaving theory behind: Why simplistic hypothesis testing is bad for International Relations,” *European Journal of International Relations* 19.3 (2013): 427-57.

Mearsheimer, John. “Why the Soviets Can’t Win Quickly in Central Europe,” *International Security* 7.1 (1982).

Meijer, Hugo. *Trading with the Enemy: The Making of US Export Control Policy Toward the People's Republic of China* (UK: Oxford University Press, 2016).

Metz, Cade. “The sadness and beauty of watching Google’s AI play Go,” *Wired*, March 11, 2016, https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/.

Michael J. Boyle, “The costs and consequences of drone warfare,” *International Affairs* 89 (2013), 27.

Michel, Arthur Holland. “Counter-Drone Systems, 2nd Edition,” *Center for the Study of the Drone*, December 2019, https://dronecenter.bard.edu/files/2019/12/CSD-CUAS-2nd-Edition-Web.pdf.

Milgrom, Paul R., and Steve Tadelis. “How Artificial Intelligence and Machine Learning Can Impact Market Design,” forthcoming in *The Economics of Artificial Intelligence*, eds. Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb (US: University of Chicago Press, 2019), 1-24. Available online: https://www.nber.org/books/agra-1.

Miller, Nicholas L. *Stopping the Bomb: The Sources of Effectiveness of US Nonproliferation Policy* (US: Cornell University Press, 2018).

Miller, Nicholas L., and Vipin Narang. “North Korea Defied the Theoretical Odds: What Can We Learn from its Successful Nuclearization?,” *Texas National Security* Review 1.2 (2018), https://tnsr.org/2018/02/north-korea-defied-theoretical-odds-can-learn-successful-nuclearization.

Mims, Christopher. “The Tiny Satellites That Will Connect Cows, Cars and Shipping Containers to the Internet,” *Wall Street Journal*, January 9, 2021, https://www.wsj.com/articles/the-tiny-satellites-that-will-connect-cows-cars-and-shipping-containers-to-the-internet-11610168400.

Mirasola, Chris. “U.S. Criticism of China’s Cybersecurity Law and the Nexus of Data Privacy and Trade Law,” *Lawfare*, October 10, 2017, https://www.lawfareblog.com/us-criticism-chinas-cybersecurity-law-and-nexus-data-privacy-and-trade-law.

Mirsky, Yisroel, and Wenke Lee. “The Creation and Detection of Deepfakes: A Survey,” *arXiv*, May 12, 2020, https://arxiv.org/pdf/2004.11138.pdf.

Mises, Ludwig von. *Human Action* (Chicago: Contemporary Books, Inc., 1963), 678-80.

Mitchell, Tom. *Machine Learning* (US: McGraw Hill, 1997).

Modelski, George, and William R. Thompson. *Leading Sectors and World Powers* (Columbia: University of South Carolina Press, 1996).

Mokyr, Joel (ed.). *The British Industrial Revolution: An Economic Perspective* (New York: Routledge, 2018).

Monteiro, Nuno. *Theory of Unipolar Politics* (New York: Cambridge University Press, 2014).

Moussseau, Michael. “The End of War: How a Robust Marketplace and Liberal Hegemony Are Leading to Perpetual World Peace,” *International Security* 44.1 (2019): 160-96.

Mozur, Paul. “One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority,” *The New York Times*, April 14, 2019, https://www.nytimes.com/2019/04/14/technology/chinasurveillance-artificial-intelligence-racial-profiling.html.

Murray, William S. “Revisiting Taiwan’s Defense Strategy,” *Naval War College Review* 61.3 (2008), https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1814&context=nwc-review.

Naim, Moises. *The End of Power: From Boardrooms to Battlefields and Churches to States, Why Being in Charge Isn't What It Used to Be* (New York: Basic Books, 2013).

Nathan, Andrew J. "China's Challenge," *Journal of Democracy* 26.1 (2015): 156-170.

Newell, Allen, and Herbert Simon. "Computer Science as Empirical Inquiry: Symbols and Search," *Communications of the ACM* 19.3 (1976): 113–26.

Newman, Jessica Cussins. "Toward AI Security: Global Aspirations for a More Resilient Future," *Center for Long-Term Cybersecurity*, 2019, https://cltc.berkeley.edu/wp-content/uploads/2019/02/Toward_AI_Security.pdf.

Ng, Andrew. "Artificial Intelligence is the New Electricity," *Stanford Graduate School of Business*, 2017, https://www.youtube.com/watch?v=21EiKfQYZXc.

Nielsen, Michael. *Neural Networks and Deep Learning* (Determination Press, 2015).

Nilsson, Nils J. *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (UK: Cambridge University Press, 2009).

Norouzzadeh, Mohammad Sadegh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune. "Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning," *PNAS* 115.25 (2018): E5716–E5725.

Norris, Guy. "USAF Unit Moves Reveal Clues To RQ-180 Ops Debut," *Aviation Week & Space Technology*, October 23, 2019, https://aviationweek.com/defense-space/usaf-unit-moves-reveal-clues-rq-180-ops-debut.

Nye, Joseph S. *The Future of Power* (New York: PublicAffairs, 2011).

O'Connor, Nuala. "Reforming the U.S. Approach to Data Protection and Privacy," *CFR*, January 30, 2018, https://www.cfr.org/report/reforming-us-approach-data-protection.

O'Kelly, Matthew, Aman Sinha, Hongseok Namkoong, John Duchi, and Russ Tedrake. "Scalable End-to-End Autonomous Vehicle Testing via Rare-event Simulation," *NeurIPS 2018*, https://arxiv.org/abs/1811.00145.

O'Sullivan, Donie. "When seeing is no longer believing: Inside the Pentagon's race against deepfake videos," *CNN*, 2019, https://www.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/.

Odgaard, Liselotte, and Thomas Galasz Nielsen. "China's Counterinsurgency Strategy in Tibet and Xinjiang," *Journal of Contemporary* China 23 (2014), 535-55.

Olah, Chris, Alexander Mordvintsev, and Ludwig Schubert. "Feature Visualization: How neural networks build up their understanding of images," *Distill*, 2017, https://distill.pub/2017/feature-visualization/. See https://distill.pub/ more generally.

Olah, Chris, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev. "The Building Blocks of Interpretability," *Distill*, https://distill.pub/2018/building-blocks/.

Olah, Chris. "Neural Networks, Manifolds, and Topology," *colah's blog*, April 6, 2014, http://colah.github.io/posts/2014-03-NN-Manifolds-Topology/.

Pape, Robert. *Bombing to Win: Air Power and Coercion in War* (US: Cornell University, 1996).

Papernot, Nicolas, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z. Berkay Celik, and Ananthram Swami. "Practical Black-Box Attacks against Machine Learning," *arXiv*, March 19, 2017, https://arxiv.org/pdf/1602.02697.pdf/.

Pauly, Reid. “Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint,” *International Security* 43.2 (2018): 151-92.

Payne, Kenneth. *Strategy, Evolution, and War: From Apes to Artificial Intelligence* (US: Georgetown University Press, 2018).

Perrault, Raymond, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles. “The AI Index 2019 Annual Report,” *Human-Centered AI Institute*, December 2019, https://hai.stanford.edu/sites/g/files/sbiybj10986/f/ai_index_2019_report.pdf.

Perry, Walter L., Brian McInnis, Carter C. Price, Susan C. Smith, and John S. Hollywood. “Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations,” *RAND*, 2013, https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf.

Peterson, Dahlia. “Designing Alternatives to China’s Repressive Surveillance State,” *CSET*, October 2020, https://cset.georgetown.edu/research/designing-alternatives-to-chinas-repressive-surveillance-state/.

Phillips, Leigh, and Michal Rozworski, *The People’s Republic of Walmart: How the World’s Biggest Corporations are Laying the Foundation for Socialism* (Verso, 2019).

Pietro, Roberto Di. “The Coming Satellite Revolution: New Business Opportunities, Scenarios, and Threats,” *Modern Diplomacy*, 2021, https://moderndiplomacy.eu/2021/05/14/the-coming-satellite-revolution-new-business-opportunities-scenarios-and-threats/.

Pike, John. “AN/APY-10 Multi-Mission Maritime and Overland Surveillance Radar,” *Global Security*, December 12, 2020, https://www.globalsecurity.org/military/systems/aircraft/systems/an-apy-10.htm.

Pike, John. “EP-3E ARIES,” *Global Security*, July 28, 2011, https://www.globalsecurity.org/intell/systems/ep-3.htm.

Pilster, Ulrich, and Tobias Bohmelt. “Coup-Proofing and Military Effectiveness in Interstate Wars, 1967–99,” *Conflict Management and Peace Science* 28, no. 4 (2011): 331-350.

Pinion, David. “The Navy and Marine Corps Need to Prepare for the Swarm of the Future,” *War on the Rocks*, March 28, 2018, https://warontherocks.com/2018/03/the-navy-and-marine-corps-must-plan-for-the-swarm-of-the-future/.

Pion-Berlin, David, and Harold Trinkunas. “Civilian Praetorianism and Military Shirking During Constitutional Crises in Latin America,” *Comparative Politics* 42, no. 4 (2010): 395-411.

Piper, Kelsey. “The case for taking AI seriously as a threat to humanity,” *Vox*, December 23, 2018, https://www.vox.com/future-perfect/2018/12/21/18126576/ai-artificial-intelligence-machine-learning-safety-alignment.

Posen, Barry. “Measuring the European Conventional Balance: Coping with Complexity in Threat Assessment,” *International Security* 9.3 (1984), 47–88.

Press, Daryl. “Simple Mathematics of Nuclear Force Analysis” (presentation, Strategic Force Analysis Bootcamp, Albuquerque, New Mexico, August 2019).

Profeta, Andrew, Andres Rodriguez, and H. Scott Clouse. “Convolutional neural networks for synthetic aperture radar classification,” paper presented at *SPIE Defense + Security*, 2016, United States, https://doi.org/10.1117/12.2225934.

Quester, George H. *Nuclear Monopoly* (UK: Routledge, 2000).

Quinlivan, James T. “Coup-proofing: Its Practice and Consequences in the Middle East,” *International Security* 24, no. 2 (1999): 131-65.

Radford, Alec, Jeffrey Wu, Dario Amodei, Daniela Amodei, Jack Clark, Miles Brundage, Ilya Sutskever, Amanda Askell, David Lansky, Danny Hernandez, and David Luan. “Better Language Models and Their Implications,” February 14, 2019, https://openai.com/blog/better-language-models/.

Radford, Alec, Jeffrey Wu, Dario Amodei, Daniela Amodei, Jack Clark, Miles Brundage, and Ilya Sutskever. “Better Language Models and Their Implications,” *OpenAI*, February 14, 2019, https://openai.com/blog/betterlanguage-models/.

Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. “Language Models are Unsupervised Multitask Learners,” *OpenAI*, February 14, 2019, https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

Rahimi, Ali and Ben Recht. “Reflections on Random Kitchen Sinks,” *arg min blog*, December 5, 2017, http://www.argmin.net/2017/12/05/kitchen-sinks/.

Rajpal, Mohit, William Blum, and Rishabh Singh, “Not all bytes are equal: Neural byte sieve for fuzzing,” *Microsoft*, 2017, https://www.microsoft.com/en-us/research/wp-content/uploads/2017/11/neural-fuzzing-mcr.pdf.

Ramesh, Aditya, Mikhail Pavlov, Gabriel Goh, Scott Gray, Mark Chen, Rewon Child, Vedant Misra, Pamela Mishkin, Gretchen Krueger, Sandhini Agarwal, and Ilya Sutskever. “DALL·E: Creating Images from Text,” *OpenAI*, 2021, https://openai.com/blog/dall-e/.

Reinsel, David, Lianfeng Wu, John F. Gantz, and John Rydning. “The China Datasphere: Primed to Be the Largest Datasphere by 2025,” *IDC*, January 2019, https://www.seagate.com/files/www-content/our-story/trends/files/data-age-china-idc.pdf.

Reisinger, Don. “A.I. Expert Says Automation Could Replace 40% of Jobs in 15 Years,” *Fortune*, http://fortune.com/2019/01/10/automation-replace-jobs/.

Remarks/Pages/Small-Satellites---Big-Data.aspx.

Resende-Santos, Joao. “Anarchy and the Emulation of Military Systems: Military Organizations and Technology in South America, 1870-1930,” *Security Studies* 5 (1996): 193-260.

Riqiang, Wu. “Certainty of Uncertainty: Nuclear Strategy with Chinese Characteristics,” *Journal of Strategic Studies* 36.4 (2013).

Riqiang, Wu. “Living with Uncertainty: Modeling China’s Nuclear Survivability,” *International Security* 44.4 (2020), 92-3.

Riqiang, Wu. “Survivability of China’s Sea-Based Nuclear Forces,” *Science and Global Security* 19.2 (2011).

Riqiang, Wu. Appendix to “Living with Uncertainty: Modeling China’s Nuclear Survivability,” *Harvard Dataverse*, 2020, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5EKNJM, 4.

Robertson, Jordan, and Michael Riley. “The Long Hack: How China Exploited a U.S. Tech Supplier,” *Bloomberg*, February 12, 2021, https://www.bloomberg.com/features/2021-supermicro/?.

Roff, Heather M. “The frame problem: The AI ‘arms race’ isn’t one,” *Bulletin of the Atomic Scientists* 75.3 (2019): 95-8.

Rosenau, William. “Special Operations Forces and Elusive Enemy Ground Targets: Lessons from Vietnam and the Persian Gulf War,” *RAND*, 2001, https://www.rand.org/pubs/monograph_reports/MR1408.html,.

Rotman, David. “AI is reinventing the way we invent,” *MIT Technology Review*, February 15, 2019, https://www.technologyreview.com/s/612898/ai-is-reinventing-the-way-we-invent/.

Rovner, Joshua. “Two kinds of catastrophe: nuclear escalation and protracted war in Asia,” *Journal of Strategic Studies* 40, no. 5 (2017): 696-730.

Rupp, Karl, and Siegfried Selberherr. “The Economic Limit to Moore’s Law,” *Proceedings of the IEEE* 98 (2010): 351-3. Available online: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5415663.

Russell, Rebecca L., Louis Kim, Lei H. Hamilton, Tomo Lazovich, Jacob A. Harer, Onur Ozdemir, Paul M. Ellingwood, and Marc W. McConley. “Automated Vulnerability Detection in Source Code Using Deep Representation Learning,” *arXiv*, November 28, 2018, https://arxiv.org/pdf/1807.04320.pdf.

Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach* (US: Prentice Hall, 2009), 5.

Russett, Bruce, and John R. Oneal. *Triangulating Peace: Democracy, Interdependence, and International Organizations* (New York: Norton, 2001).

Russett, Bruce. *Grasping the Democratic Peace: Principles for a Post–Cold War World* (New Jersey: Princeton University Press, 1993).

Saalman, Lora (ed.). *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume II: East Asian Perspectives* (SIPRI: October 2019). Available online: https://www.sipri.org/publications/2019/other-publications/impact-artificial-intelligence-strategic-stability-and-nuclear-risk-volume-ii-east-asian.

Saalman, Lora. “Fear of false negatives: AI and China’s nuclear posture,” *Bulletin of the Atomic Scientists*, April 24, 2018, https://thebulletin.org/2018/04/fear-of-false-negatives-ai-and-chinas-nuclear-posture/.

Sacks, Samm, and Lorand Laskai. “China’s Privacy Conundrum,” *Slate*, February 7, 2019, https://slate.com/technology/2019/02/china-consumer-data-protection-privacy-surveillance.html.

Sankaran, Jaganath. “A Different Use for Artificial Intelligence in Nuclear Weapons Command and Control,” *War on the Rocks*, April 25, 2019, https://warontherocks.com/2019/04/a-different-use-for-artificial-intelligence-in-nuclear-weapons-command-and-control/.

Sayler, Kelley M. “Artificial Intelligence and National Security,” *Congressional Research Service*, 2020, https://fas.org/sgp/crs/natsec/R45178.pdf, 34.

Sayler, Kelley M., “Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems,” *CRS*, December 19, 2019, https://crsreports.congress.gov/product/pdf/IF/IF11150.

Scharre, Paul, and Michael C. Horowitz. “An Introduction to Autonomy in Weapon Systems,” *CNAS*, 2015, https://www.cnas.org/publications/reports/an-introduction-to-autonomy-in-weapon-systems.

Scharre, Paul. “Counter-Swarm: A Guide to Defeating Robotic Swarms,” *War on the Rocks*, March 31, 2015, https://warontherocks.com/2015/03/counter-swarm-a-guide-to-defeating-robotic-swarms/.

Scharre, Paul. “Robotics on the Battlefield, Part I: Range, Persistence, and Daring,” *CNAS*, 2014.

Scharre, Paul. “Robotics on the Battlefield, Part II: The Coming Swarm,” *CNAS*, October 2014, https://s3.amazonaws.com/files.cnas.org/documents/CNAS_TheComingSwarm_Scharre.pdf?

Scharre, Paul. “Robots at War and the Quality of Quantity,” *War on the Rocks*, February 26, 2015, https://warontherocks.com/2015/02/robots-at-war-and-the-quality-of-quantity/.

Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War* (New York: W. W. Norton and Company, 2018).

Schelling, Thomas C. *Arms and Influence* (New Haven: Yale University Press, 2008).

Schmidt, Eric, Robert O. Work, Safra Catz, Steve Chie, Mignon Clyburn, Christopher Darby, Kenneth Ford, Jose-Marie Griffiths, Eric Horvitz, Andrew Jassy, Gilman Louie, William Mark, Jason Matheny, Katharina McFarland, and Andrew Moore. “Interim Report,” *National Security Commission on Artificial Intelligence*, November 2019, 7. Available online: https://www.epic.org/foia/epic-v-ai-commission/AI-Commission-Interim-Report-Nov-2019.pdf.

Schmidt, Eric, Robert Work, Safra Catz, Eric Horvitz, Steve Chien, Andrew Jassy, Mignon Clyburn, Gilman Louie, Chris Darby, William Mark, Kenneth Ford, Jason Matheny, Jose-Marie Griffiths, Katharina McFarland, and Andrew Moore, “Final Report,” *National Security Commission on Artificial Intelligence*, March 2021, https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf, 1, 22-3.

Schneider, Jacquelyn. “Cyber and Crisis Escalation: Insights from Wargaming,” *USASOC Futures Forum*, 2017.

Schneider, Jacquelyn. “The capability/vulnerability paradox and military revolutions: Implications for computing, cyber, and the onset of war,” *Journal of Strategic Studies* 42.6 (2019): 841-63.

Schneider, Todd, Gee Hee Hong, and Anh Van Le. “Land of the Rising Robots,” *IMF*, June 2018, https://www.imf.org/external/pubs/ft/fandd/2018/06/japan-labor-force-artificial-intelligence-and-robots/schneider.pdf.

Schuett, Jonas. “A Legal Definition of AI,” *arXiv*, September 4, 2019, https://arxiv.org/pdf/1909.01095.pdf.

Schwegmann, C. P., W. Kleynhans, and B. P. Salmon. “The development of deep learning in synthetic aperture radar imagery,” paper presented at the *International Workshop on Remote Sensing with Intelligent Processing* (RSIP), 2017, China, https://doi.org/10.1109/RSIP.2017.7958802.

Schwegmann, C.P., W. Kleynhans, B.P. Salmon, L.W. Mdakane, and R.G.V. Meyer, “Very deep learning for ship discrimination in Synthetic Aperture Radar imagery,” paper presented at the *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), China, 2016, https://ieeexplore.ieee.org/abstract/document/7729017/.

Scoblic, J. Peter, and Philip E. Tetlock. “A Better Crystal Ball: The Right Way to Think About the Future,” *Foreign Affairs*, 2020, https://www.foreignaffairs.com/articles/united-states/2020-10-13/better-crystal-ball.

Sculley, D., Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, and Michael Young. “Machine Learning: The High Interest Credit Card of Technical Debt,” *Google*, 2014, https://research.google/pubs/pub43146/.

Sechser, Todd S., Neil Narang, and Caitlin Talmadge. “Emerging technologies and strategic stability in peacetime, crisis, and war,” *Journal of Strategic Studies* 42.6 (2019): 727-35.

Senior, Andrew W., Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Zidek, Alexander W. R. Nelson, Alex Bridgland, Hugo Penedones, Stig Petersen, Karen Simonyan, Steve Crossan, Pushmeet Kohli, David T. Jones, David Silver, Koray Kavukcuoglu, and Demis Hassabis. “Improved protein structure using potentials from deep learning,” *Nature* (2020), https://www.nature.com/articles/s41586-019-1923-7.

Sepp, Eric M. “Deeply Buried Facilities: Implications for Military Operations,” *Occasional Paper* 14 (2000), 25, 31.

Shachtman, Noah. “This Rock Could Spy on You for Decades,” *WIRED*, May 29, 2012, https://www.wired.com/2012/05/spy-rock/.

Shafahi, Ali, W. Ronny Huang, Mahyar Najibi, Octavian Suciu, Christoph Studer, Tudor Dumitras, and Tom Goldstein. “Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks,” *NeurIPS*, 2018, https://papers.nips.cc/paper/7849-poison-frogs-targeted-clean-label-poisoning-attacks-on-neural-networks.pdf.

Shambaugh, David. “The Coming Chinese Crackup,” *The Wall Street Journal*, March 6, 2015, https://www.wsj.com/articles/the-coming-chinese-crack-up-1425659198.

Shane, Scott, and David E. Sanger. “Drone Crash in Iran Reveals Secret U.S. Surveillance Effort,” *The New York Times*, December 7, 2011, https://www.nytimes.com/2011/12/08/world/middleeast/drone-crash-in-iran-reveals-secret-us-surveillance-bid.html.

Shlapak, David A., David T. Orletsky, Toy I. Reid, Murray Scot Tanner, and Barry Wilson. “A Question of Balance: Political Context and Military Aspects of the China-Taiwan Dispute,” *RAND*, 2009, https://www.rand.org/pubs/monographs/MG888.html.

Shmuel, Shmuel. “The Coming Swarm Might Be Dead on Arrival,” *War on the Rocks*, September 10, 2018, https://warontherocks.com/2018/09/the-coming-swarm-might-be-dead-on-arrival/.

Shortliffe, Edward H. “Medical Expert Systems – Knowledge Tools for Physicians,” *Western Journal of Medicine* 145.6 (1986): 830-9.

Shugart, Thomas, and Javier Gonzalez. “First Strike: China’s Missile Threat to U.S. Bases in Asia,” *CNAS*, June 2017, https://s3.amazonaws.com/files.cnas.org/documents/CNASReport-FirstStrike-Final.pdf?.

Silver, David, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis, “Mastering the game of Go without human knowledge.” *Nature* 550 (2017): 354–59.

Silver, David, Thomas Hubert, Julian Schrittwieser, and Demis Hassabis, “AlphaZero: Shedding new light on chess, shogi, and Go,” *DeepMind*, December 6, 2018, https://deepmind.com/blog/article/alphazero-shedding-new-light-grand-games-chess-shogi-and-go.

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, and Demis Hassabis. “Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm,” *arXiv*, December 2017, https://arxiv.org/pdf/1712.01815.pdf.

Skitka, Linda J., Kathleen L. Mosier, and Mark Burdick. “Does automation bias decision-making?”, *International Journal of Human-Computer Studies* 51 (1999): 991-1006.

Solaiman, Irene, Miles Brundage, Jack Clark, Amanda Askell, Ariel Herbert-Voss, Jeff Wu, Alec Radford, Gretchen Krueger, Jong Wook Kim, Sarah Kreps, Miles McCain, Alex Newhouse, Jason Blazakis, Kris McGuffie, and Jasmine Wang. “Release Strategies and the Social Impacts of Language Models,” *arXiv*, November 13, 2019, https://arxiv.org/ftp/arxiv/papers/1908/1908.09203.pdf.

Soldin, Ryan J., Douglas N. MacDonald, Matthew Reisman, Latisha R. Konz, Roger Rouse, and Timothy L. Overman. “HySARNet: a hybrid machine learning approach to synthetic aperture radar automatic target recognition,” paper presented at *SPIE Defense + Commercial Sensing*, 2019, United States, https://doi.org/10.1117/12.2518155

Somers, James. “Is AI Riding a One-Trick Pony?”, *MIT Technology Review*, September 29, 2017, https://www.technologyreview.com/s/608911/is-ai-riding-a-one-trick-pony/.

Spencer, John. “The City Is Not Neutral: Why Urban Warfare Is So Hard,” *Modern War Institute*, March 4, 2020, https://mwi.usma.edu/city-not-neutral-urban-warfare-hard/.

Stefanick, Tom. “AI in the Aether: Military Information Conflict,” in *The Global Race for Technological Superiority*, ed. Fabio Rugge (Brookings, 2019), https://www.brookings.edu/wp-content/uploads/2019/12/FP_20191211_military_information_conflict_stefanick.pdf, 112-30.

Steinhardt, Jacob, Pang Wei Koh, and Percy Liang. “Certified Defenses for Data Poisoning Attacks,” NIPS, 2017, https://papers.nips.cc/paper/6943-certified-defenses-for-data-poisoning-attacks.pdf.

Stewart, Phil. “Deep in the Pentagon, a secret AI program to find hidden nuclear missiles,” *Reuters*, June 5, 2018, https://www.reuters.com/article/us-usa-pentagon-missiles-ai-insight/deep-in-the-pentagon-a-secret-ai-program-to-find-hidden-nuclear-missiles-idUSKCN1J114J.

Stillion, John, and Scott Perdue. “Air Combat Past, Present, and Future,” *RAND*, August 2008, https://www.defenseindustrydaily.com/files/2008_RAND_Pacific_View_Air_Combat_Briefing.pdf.

Stokes, Jon. “Why a Chinese invasion of Taiwan would be a catastrophe for China and the world,” *Doxa*, 2021, https://doxa.substack.com/p/why-a-chinese-invasion-of-taiwan.

Stokes, Mark A. “China's Nuclear Warheads Storage and Handling System,” *Project 2049 Institute*, March 12, 2010, https://project2049.net/wp-content/uploads/2018/05/chinas_nuclear_warhead_storage_and_handling_system.pdf.

Stubberud, Stephen C., and Kathleen A. Kramer. “UAVs working in conjunction with unattended ground sensors,” *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*.

Su, Alice. “Can fact-checkers save Taiwan from a flood of Chinese fake news?,” *Los Angeles Times*, December 16, 2019, https://www.latimes.com/world-nation/story/2019-12-16/taiwan-the-new-frontier-of-disinformation-battles-chinese-fake-news-as-elections-approach.

Sullivan, Brian. “Close encounters of the fishy kind,” *Google*, June 8, 2018, https://www.blog.google/products/earth/close-encounters-fishy-kind/.

Sun, Leo. “Why Intel's Foundry Plans Don't Make Any Sense,” *The Motley Fool*, 2021, https://www.fool.com/investing/2021/05/03/why-intel-foundry-plans-dont-make-any-sense/.

Sutton, Rich. “The Bitter Lesson,” *Incomplete Ideas*, March 13, 2019, http://www.incompleteideas.net/IncIdeas/BitterLesson.html.

Talmadge, Caitlin. “Would China Go Nuclear? Assessing the Risk of Chinese Nuclear Escalation in a Conventional War with the United States,” *International Security* 41, no. 4 (2017): 50-92.

Talmadge, Caitlin. *The Dictator’s Army: Battlefield Effectiveness in Authoritarian Regimes* (Ithaca: Cornell University Press, 2015).

Tarraf, Danielle C., William Shelton, Edward Parker, Brien Alkire, Diana Gehlhaus Carew, Justin Grana, Alexis Levehahl, Jasmin Leveille, Jared Mondschein, James Ryseff, Ali Wyne, Dan Elinoff, Edward Geist, Benjamin N. Harris, Eric Hui, Cedric Kenney, Aydne Newberry, Chandler Sachs, Peter Schirmer, Danielle Schlang, Vicotria Smith, Abbie Tingstad, Padmaja Vedula, and Kristin Warren. “The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations,” *RAND*, 2019, https://www.rand.org/pubs/research_reports/RR4229.html.

Tetlock, Philip E., and Dan Gardner. *Superforecasting: The Art and Science of Prediction* (NY: Broadway Books, 2016).

Thomas, Jim, John Stillion, and Iskander Rehman. "Hard ROC 2.0: Taiwan and Deterrence through Protraction," *CSBA*, 2014, https://csbaonline.org/uploads/documents/2014-10-01_CSBA-TaiwanReport-1.pdf.

Thompson, Neil C., and Svenja Spanuth. "The Decline of Computers as a General Purpose Technology: Why Deep Learning and the End of Moore's Law are Fragmenting Computing," *SSRN*, 2018, http://ide.mit.edu/sites/default/files/publications/SSRN-id3287769.pdf.

Ting-Fang, Cheng, and Lauly Li. "The great US-China tech decoupling: Where are we now?," *Nikkei Asia*, 2019, asia.nikkei.com/Economy/Trade-war/The-great-US-China-tech-decoupling-Where-are-we-now.

Toonders, Joris. "Data Is the New Oil of the Digital Economy," *WIRED*, 2014, https://www.wired.com/insights/2014/07/data-new-oil-digital-economy/.

Topychkanov, Petr (ed.). *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume III: South Asian Perspectives* (SIPRI: April 2020). Available online: https://www.sipri.org/publications/2020/other-publications/impact-artificial-intelligence-strategic-stability-and-nuclear-risk-volume-iii-south-asian.

Tucker, Patrick. "How AI Will Transform Anti-Submarine Warfare," *Defense One*, July 1, 2019, https://www.defenseone.com/technology/2019/07/how-ai-will-transform-anti-submarine-warfare/158121/.

Turek, Matt. "Explainable Artificial Intelligence (XAI)," *DARPA*, 2018, https://www.darpa.mil/program/explainable-artificial-intelligence.

Turner, Jennifer M. "The Cost of Credible Deterrence in Taiwan," *War on the Rocks*, January 13, 2016, https://warontherocks.com/2016/01/the-cost-of-credible-deterrence-in-taiwan/.

Unver, Akin H. "Artificial Intelligence, Authoritarianism and the Future of Political Systems," *Center for Economics and Foreign Policy Studies*, 2018.

Valentino, Benjamin, Paul Huth, and Dylan Balch-Lindsay. "'Draining the Sea': Mass Killing and Guerrilla Warfare," *International Organization* 58.2 (2004): 375-407.

Van Horn, Grant, and Pietro Perona. "The Devil is in the Tails: Fine-grained Classification in the Wild," *arXiv*, September 5, 2017, https://arxiv.org/pdf/1709.01450.pdf.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. "Attention is All You Need," *arXiv*, December 16, 2017, https://arxiv.org/pdf/1706.03762.pdf.

Vick, Alan J., Richard M. Moore, Bruce R. Pirnie, and John Stillion. "Aerospace Operations Against Elusive Ground Targets," *RAND*, 2001, https://www.rand.org/pubs/monograph_reports/MR1398.html.

Vinyals, Oriol, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature* 575 (2019).

Waltz, Kenneth. *Theory of International Politics* (New York: McGraw-Hill, 1979).

Wang, Maya. ""Eradicating Ideological Viruses": China's Campaign of Repression Against Xinjiang's Muslims," *Human Rights Watch*, September 9, 2018, https://www.hrw.org/report/2018/09/09/eradicating-ideological-viruses/chinas-campaign-repression-against-xinjiangs.

Wang, Maya. "China's Algorithms of Repression: Reverse Engineering a Xinjiang Police Mass Surveillance App," *Human Rights Watch*, May 1, 2019, https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverseengineering-xinjiang-police-mass-surveillance.

Wang, Yuhua, and Carl Minzner. "The Rise of the Chinese Security State," *The China Quarterly* 222 (2015): 339-59.

Wang, Yuhua. "Coercive capacity and the durability of the Chinese communist state," *Communist and Post-Communist Studies* 47 (2014): 13-25.

Wang, Yuhua. "Empowering the Police: How the Chinese Communist Party Manages Its Coercive Leaders," *The China Quarterly* 219 (2014): 625-48.

Wang, Yu-Xiong, Deva Ramanan, and Marital Hebert. "Learning to Model the Tail," paper presented at the *31st Conference on Neural Information Processing Systems*, Long Beach, California, 2017. Available online: https://papers.nips.cc/paper/7278-learning-to-model-the-tail.pdf.

Watts, Barry D., and Thomas A. Keaney. "Effects and Effectiveness," in *Gulf War Air Power Survey, Volume II: Operations and Effects and Effectiveness*, ed. Eliot A. Cohen (Washington, DC: US Government Publishing Office, 1993), 330-40. Available online: https://apps.dtic.mil/dtic/tr/fulltext/u2/a279742.pdf.

Weinberg, Justin. "Philosophers On GPT-3 (updated with replies by GPT-3)," *Daily Nous*, 2020, https://dailynous.com/2020/07/30/philosophers-gpt-3/therapist.

Weiner, Robert J. "Is the World Oil Market 'One Great Pool?'" *Energy Journal* 12.3 (1991), 95-108.

Weisgerber, Marcus. "The Increasingly Automated Hunt for Mobile Missile Launchers," *Defense One*, April 26, 2016, https://www.defenseone.com/technology/2016/04/increasingly-automated-hunt-mobile-missile-launchers/127864/.

Wellerstein, Alex. "The trouble with airbursts," *Restricted Data*, December 6, 2013, http://blog.nuclearsecrecy.com/2013/12/06/trouble-airbursts/.

Wiener, Norbert. "Some Moral and Technical Consequences of Automation," *Science* 131 (1960), 1358.

Winick, Erin. "Every study we could find on what automation will do to jobs, in one chart," *MIT Technology Review*, January 25, 2018, https://www.technologyreview.com/s/610005/every-study-we-could-find-on what-automation-will-do-to-jobs-in-one-chart/.

Wohlforth, William C., Dmitry Adamsky, Theo Farrell, Adam Grissom, Thomas G. Mahnken, and Michael C. Horowitz. "H-Diplo/ISSF Roundtable Review of Michael C. Horowitz. The Diffusion of Military Power: Causes and Consequences for International Relations (2010)," *Roundtable* 3.10 (2012), https://issforum.org/ISSF/PDF/ISSF-Roundtable-3-10.pdf.

Wong, Samantha. "Road density in China 2008-2019," *Statista*, November 24, 2020, https://www.statista.com/statistics/258345/road-density-in-china/.

Wong, Yuna Huh, John Yurchak, Robert W. Button, Aaron Frank, Burgess Laird, Osonde A. Osoba, Randall Steeb, Benjamin N. Harris, and Sebastian Joon Bae. "Deterrence in the Age of Thinking Machines," *RAND*, 2020, https://www.rand.org/pubs/research_reports/RR2797.html.

Woodward, C. Vann. “The Age of Reinterpretation,” *The American Historical Review* 66 (1960): 1-19.

Work, Robert O., and Greg Grant. “Beating the Americans at Their Own Game: An Offset Strategy with Chinese Characteristics,” *CNAS*, June 6, 2019, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Report-Work-Offset-final-B.pdf?

Work, Robert O., and Shawn Brimley. “20YY: Preparing for War in the Robotic Age,” *CNAS*, January 2014, https://s3.amazonaws.com/files.cnas.org/documents/CNAS_20YY_WorkBrimley.pdf?.

Wright, Nicholas. “How Artificial Intelligence Will Reshape the Global Order,” *Foreign Affairs*, July 10, 2018, https://www.foreignaffairs.com/articles/world/2018-07-10/howartificial-intelligence-will-reshape-global-order.

Wu, Yonghui, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, and Jeffrey Dean, “Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation,” *arXiv*, October 6, 2016, https://arxiv.org/pdf/1609.08144.pdf.

Xiaomeng, Lu, Li Manyi, and Samm Sacks. “What the Facebook Scandal Means in a Land without Facebook: A Look at China’s Burgeoning Data Protection Regime,” *CSIS*, April 25, 2018, https://www.csis.org/analysis/what-facebook-scandal-means-land-without-facebook-look-chinas-burgeoning-data-protection.

Xiaoning, Zhu. “Analysis of military application of UAV swarm technology,” paper presented at the *3rd International Conference on Unmanned Systems* (November 27-28, 2020, Harbin, China), https://ieeexplore.ieee.org/abstract/document/9274974.

Xu, Beina, and Eleanor Albert. “Media Censorship in China,” *Council on Foreign Relations*, February 17, 2017, https://www.cfr.org/backgrounder/media-censorship-china.

Yang, Bo, and Min Liu. “Keeping in Touch with Collaborative UAVs: A Deep Reinforcement Learning Approach,” *IJCAI-18*, 2018, https://www.ijcai.org/Proceedings/2018/0078.pdf.

Ying, Fu (trans. Brian Tse and Jeffrey Ding). “A Preliminary Analysis of the Impact of AI on International Relations,” *Quarterly Journal of International Politics* 4 (2019): 1-18. Translation available at https://chinai.substack.com/p/chinai-67-fu-ying-on-ai-the-international.

Zegart, Amy. “Cheap fights, credible threats: The future of armed drones and coercion,” *Journal of Security Studies* 32.1 (2020): 6-46.

Zellers, Rowan, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. “Defending Against Neural Fake News,” *arXiv*, October 29, 2019, https://arxiv.org/pdf/1905.12616.pdf.

Zeng, Jinghan. “Artificial intelligence and China's authoritarian governance,” *International Affairs* 96.6 (2020): 1441-59.

Zenz, Adrian. “China’s Domestic Security Spending: An Analysis of Available Data,” *China Brief* 18.4 (2018), https://jamestown.org/program/chinas-domestic-security-spending-analysis-available-data/.

Zhang, Baobao, and Allan Dafoe. “Artificial Intelligence: American Attitudes and Trends,” *Centre for the Governance of AI*, 2019.

Zhang, Baobao, and Allan Dafoe. “U.S. Public Opinion on the Governance of Artificial Intelligence,” *Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society*, 2020.

Zhang, Baobao, Markus Anderljung, Lauren Kahn, Noemi Dreksler, Michael C. Horowitz, and Allan Dafoe. “Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers,” *Center for the Governance of AI*, 2019.

Zhao, Tong. “Tides of Change: China's Nuclear Ballistic Missile Submarines and Strategic Stability,” *Carnegie Endowment for International Peace*, 2018, https://carnegieendowment.org/files/Zhao_SSBN_final.pdf.

Zisk, Kimberly. *Engaging the Enemy: Organization Theory and Soviet Military Innovation, 1955-1991* (Princeton: Princeton University Press, 1993).

Zwetsloot, Remco, Helen Toner, and Jeffrey Ding, “Beyond the AI Arms Race: America, China, and the Dangers of Zero-Sum Thinking,” *Foreign Affairs*, November 16, 2018, https://www.foreignaffairs.com/reviews/review-essay/2018-11-16/beyond-ai-arms-race.

Zwetsloot, Remco, James Dunham, Zachary Arnold, and Tina Huang, “Keeping Top AI Talent in the United States: Finding and Policy Options for International Graduate Student.” *CSET*, December 2019, https://cset.georgetown.edu/wp-content/uploads/Keeping-Top-AI-Talent-in-the-United-States.pdf.