Founders Pledge

Centre for International Governance Innovation

Great Power Competition and Transformative Technologies

STEPHEN CLARE and CHRISTIAN RUHL

January 2024

Table of Contents

Disclaimers and acknowledgments

With thanks to Tom Barnes, Haydn Belfield, Rosella Cappella Zielinski, Matt Lerner, Andrew Reddie, Paul Scharre, and Jim Scouras, for their advice and assistance with this project. GPT4 assisted in the brainstorming, outlining, and editing of this report.

Co-published with the Centre for International Governance Innovation (CIGI)

CIGI is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

Introduction

Many interventions to reduce global catastrophic risks (GCR) are cause-specific: fund AI safety research, fund pandemic-proof PPE, study the effects of specific future weapon systems, and so on.[1] This report attempts to step back and look at a root driver of catastrophic risk: technology competition between great power states. Such competition drives governments to race for military superiority, become paranoid about their rivals' intentions, and build secretive bureaucracies. Each of these in turn may raise existential risk.[2]

Some analysts have claimed that "most of the risk we face [i.e., humanity faces] comes from scenarios where there is a hot or cold war between great powers" leading to global catastrophes involving emerging technologies.[3] This report investigates this claim by examining how **strategic competition between great powers over transformative technologies** affects various GCRs**.** We focus specifically on the worst possible outcomes of such competition — global catastrophes and existential threats — and seek to provide a framework for understanding these risks. We also assess the importance, neglectedness, and tractability of these issues and suggest what, if anything, philanthropists could do to mitigate them.

In this report, we argue that:

1. Strategic competition has historically accelerated and even caused the development and deployment of transformative technologies.
2. Such competition may affect catastrophic risks via multiple pathways. We investigate seven specific effects, grouped into three categories: raising the risk of war, raising the risk of accidents, and causing dangerous technologies to proliferate.
3. There is a lot of uncertainty about the magnitude and, in some cases, the sign of these effects, though in expectation they seem quite important.
4. This cause appears to receive little targeted philanthropic investment.
5. There are some plausible options for philanthropic intervention, such as funding research to improve the culture of safety in government bureaucracies.
6. But tractability remains the biggest concern — states have strong incentives to compete, and actors within states have incentives to encourage this competition.

The report is structured as follows:

1. **What is the Problem?**
   We define key concepts, provide an overview of technology competition over transformative

[1] For a list of GCR mitigation interventions, see GCR Policy, "Policy Ideas," https://www.gcrpolicy.com/ideas.
[2] Richard Danzig, "Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority" (CNAS, 2018), https://www.cnas.org/publications/reports/technology-roulette.
[3] William MacAskill, *What We Owe the Future* (New York: Basic Books, 2022), 274.

technologies, and introduce a framework for thinking about the risks of this competition via our three main pathways.

2. **Importance**
   Next, we argue that modern strategic competition will influence the development of multiple transformative technologies, each of which could influence humanity's future trajectory. We then examine each of the risk pathways outlined above, and consider whether ongoing competition is likely to increase or decrease risk. For most of the pathways, **more intense strategic competition seems likely to raise catastrophic risk**.

3. **Neglectedness**
   This short section provides a rough attempt to understand the neglectedness of this problem. This is a challenging question, given that national competitiveness is a major focus for many philanthropists, think tanks, and government offices. We argue, though, that its intersection with extreme risks receives relatively little attention.

4. **Tractability**
   This section briefly touches on the problem of tractability. We argue that while affecting geopolitical outcomes is clearly challenging, **progress on specific issues seems possible**.

   a. **Interventions**
      Following from the tractability discussion, we briefly sketch possible interventions in our model, including non-proliferation measures, improvements in threat assessment, policy advocacy, and more.

5. **Conclusion and Next Steps**

This report is **part of an evolving conversation on existential risks and great power conflict** rather than a definitive treatment of the subject. Broadly we hope to strengthen the claim that competition between the world's most well-resourced and powerful states is likely one of the biggest drivers of existential risk. We also hope to emphasize that some opportunities for philanthropists or individuals to mitigate the risks of this competition exist, and that, because this root driver of risk is *leveraged,* these possibilities warrant further exploration.

What is the Problem?

This section sketches an outline of the problem of transformative technology competition between major militaries. Competition over high-risk technologies drove some anthropogenic catastrophic risk during the twentieth century and continues to shape the threat landscape today; it shapes the risk of war and the development and deployment of transformative technologies like synthetic

biology and artificial intelligence. After this discussion, the next section (Importance) assesses the scale of this problem.

Definitions and Key Concepts

In this report, the term **transformative technologies** refers to technologies that could significantly alter the trajectory of human civilization. These include new general-purpose technologies (GPTs) with large and far-reaching societal impact, like electricity and, potentially, artificial intelligence.[4] Transformative technologies also include technologies such as new weapon systems that could either dramatically alter the global balance of power or pose an **existential risk**, threatening to harm humanity's long-term future.

In this report, we discuss transformative technologies whose development is influenced by international military competition and which could plausibly be developed in roughly the next century. These include:

- AI and AI-enabled weapons (like autonomous weapons)
- Synthetic biology, biodefense, and biological weapons
- Next-generation nuclear weapons and delivery systems
- Other currently-unknown technologies whose development is hard to predict

We focus our analysis on **great power states**. Great powers are states with global interests and the means to defend them, including military capabilities that enable them to fight major wars.[5] We focus on these countries because *how* their governments compete, cooperate, and leverage the enormous resources they control can strongly influence how transformative technologies develop. In practice, this report focuses on the United States and China, and to a lesser extent Russia (as a nuclear-armed state that shares some but not all the attributes of a great power) and India (as a potential rising power).

Technological superiority has become a key goal of modern militaries. As a result, great powers are likely to engage in **transformative technology competition**. Such competition involves racing to develop such technologies first and control how they are deployed and governed. It can include both trying to speed up domestic technology development, such as by subsidizing companies and investing in research efforts, and trying to slow down competitors by, for example, setting up trade policies that seek to deny other actors access to useful technologies and knowledge.

Developing the technologies over which great powers compete often comes with considerable risks. Many military technologies are intrinsically destructive. Moreover, their development is shrouded in secrecy and under pressure to proceed quickly. And new weapons are difficult to test and often

---

[4] For a discussion of the military effects of GPT revolutions, see Jeffrey Ding and Allan Dafoe, "Engines of Power: Electricity, AI, and General-Purpose, Military Transformations," *European Journal of International Security*, February 7, 2023, 1–18, https://doi.org/10.1017/eis.2023.1.

[5] For a discussion of the difficulty of defining "great powers," see Clare, *Great Power Conflict*, 20.

deployed in unfamiliar, high-stress environments. In the near future, for example, we may see powerful AI-enabled systems slip out of human control and cause enormous damage. Engineered pathogens could leak out of laboratories and infect and kill billions. And other future technologies and weapon systems could pose similar risks that are simply impossible to predict today.

Because these military technologies are inherently destructive, Richard Danzig, a former US Secretary of the Navy, has called these races "**technology roulette**." Pursuing military superiority through technological development, Danzig writes, is like "loading increasing numbers of bullets into increasing numbers of revolvers held to the head of humanity."[6] This report explicitly seeks to build on Danzig's work by developing a more detailed model of how international competition affects technological development and identifying what private philanthropists can do to mitigate these risks.

Importantly, these risks go beyond great power war per se; even well-managed great power competition that does not go into war, but finds other outlets for strategic competition, could be a major threat in this model.

By focusing on high-risk technology competition *in general*, this report is taking a threat-agnostic approach to existential risk factors. We discuss individual risks previous analyses have often focused on — AI, synthetic biology, etc. — but also attempt to look at the competitive dynamics that drive the emergence of these threats in the first place. In this sense, great power competition is an **existential risk factor**: a phenomenon that does not directly threaten humanity's long-term future, but which increases the overall level of risk we face. Stepping back thus works as an **impact multiplier**. By focusing our efforts earlier in the causal chain, we may help reduce the threat posed by multiple anthropogenic risks at once.

Overview of Technology Competition

Why should people concerned about major technological risks pay particular attention to international technology competition? In short:

1. **Technological superiority is a core goal of modern states.** States especially prize technologies with potential military applications
2. **Many of these technologies are inherently destructive** or destabilizing
3. **Intensifying technological competition therefore may speed up the development of destructive technologies**
4. Additionally, intensifying competition may lead to a **race to the bottom on safety**, where testing and evaluation takes a backseat to the pursuit of technological superiority

[6] Richard Danzig, "Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority" (Center for a New American Security, 2018), https://www.cnas.org/publications/reports/technology-roulette.

The first point is that technological superiority is a *core goal* of leading modern militaries.[7]

Broadly, the pursuit of technological superiority has been at the center of U.S. military thinking and its three "offsets strategies" since the early Cold War:

1. **First Offset** — The Eisenhower administration's decision to focus on nuclear capabilities rather than attempting to match the USSR "man-for-man and tank-for-tank"[8]
2. **Second Offset** — Confronted by nuclear parity with the Soviet Union, in the 1970s, defense strategists decided to instead focus on emerging technologies that enabled precision-guided munitions, stealth technologies, the GPS, computer networks and other systems behind contemporary high-tech war[9]
3. **Third Offset** — Since 2014, the U.S. Department of Defense has been operating with the assumption that another "offset" is necessary and that "the technological sauce of the Third Offset is going to be advances in Artificial Intelligence (AI) and autonomy," in the words of Deputy Secretary Robert Work, who helped design the strategy[10]

The demonstration of the success of the Second Offset Strategy during the First Gulf War, combined with the humiliation of the Third Taiwan Strait Crisis, arguably shaped China's pursuit since the 1990s of technological parity with, then superiority over, the United States.[11] The belief in the importance of high-tech warfare is summed up in Xi Jingping's statement that "advanced technology is the sharp weapon of the modern state."[12]

Militaries also do not face the same constraints and incentives that private actors do. As Danzig writes, "Richly endowed nation-state rivals operate with great paranoia and little inhibition."[13] Most importantly (and most obviously), militaries can develop and use weapons and weapon systems that can be highly destructive, like nuclear weapons and (at some points in history) biological weapons.

Technology competition can vary in intensity:

---

[7] This is discussed in detail with reference to strategic competition over computer chips in Chris Miller, *Chip War: The Fight for the World's Most Critical Technology*, 2022.
[8] Robert O Work and Greg Grant, "Beating the Americans at Their Own Game," *CNAS*, 2019, https://www.cnas.org/publications/reports/beating-the-americans-at-their-own-game.
[9] Ibid.
[10] "Remarks by Deputy Secretary Work on Third Offset Strategy," U.S. Department of Defense, accessed January 12, 2024, https://www.defense.gov/News/Speeches/Speech/Article/753482/remarks-by-deputy-secretary-work-on-third-offset-strategy/https%3A%2F%2Fwww.defense.gov%2FNews%2FSpeeches%2FSpeech%2FArticle%2F753482%2Fremarks-by-deputy-secretary-work-on-third-offset-strategy%2F.
[11] For a full discussion of this claim, see Work and Grant, *Beating the Americans at their Own Game* https://www.cnas.org/publications/reports/beating-the-americans-at-their-own-game.
[12] Ibid.
[13] Richard Danzig, "Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority" (CNAS, 2018), 8, https://www.cnas.org/publications/reports/technology-roulette.

- Intense competition can drive an all-out arms race, such as the nuclear arms race between the US and USSR between 1945 and 1990
- Mild competition can enable international cooperation, such as internationally agreed and enforced bans on certain technologies like human cloning
- A moderate degree of competition can lead to states cooperating in some domains and competing in others; for example, agreeing to international controls on nuclear materials to limit proliferation while maintaining domestic control of their own nuclear materials

We hypothesize that the intensity of technology competition is driven by multiple factors, including how much trust or suspicion key actors view each other with, the current balance of power, the feasibility of monitoring and enforcement of international agreements, and the perceived stakes of the competition. Furthermore, we hypothesize that more intense competitions are more likely to raise global risks.

This report seeks to investigate these dynamics in greater detail and make some progress towards estimating their strength.

Publicly-available U.S. Strategy documents emphasize the importance of technology competition and the various levers available to states to pursue such competition. The U.S. 2022 National Security Strategy (NSS) emphasizes "prioritize maintaining an enduring competitive edge over the PRC [People's Republic of China]" and that technology competition will be a major pillar in this strategy.[14] To maintain this competitive edge, the Strategy points to:

- **Domestic Investments**: "The CHIPS and Science Act authorizes $280 billion for civilian investment in research and development, especially in critical sectors such as semiconductors and advanced computing, next-generation communications, clean energy technologies, and biotechnologies" (15) and "Through the National Biotechnology and Biomanufacturing Initiative, we are investing more than $2 billion to harness the full potential of biotechnology and biomanufacturing" (15)
- **Market Shaping and Industrial Policy**: "We are identifying and investing in key areas where private industry, on its own, has not mobilized to protect our core economic and national security interests, including bolstering our national resilience" (14)
- **Supply-Chain Shaping**: "We recognize the importance of the semiconductor supply chain to our competitiveness and our national security, and we are seeking to reinvigorate the semiconductor industry in the United States" (15)
- **Immigration Policy**: "we will also continue to make America the destination of choice for talent around the world [...] And we will take further measures to ensure the United States remains the world's top destination for talent" (15-16)

[14] The White House, "The Biden-Harris Administration's National Security Strategy," The White House, October 12, 2022, https://www.whitehouse.gov/briefing-room/statements-releases/2022/10/12/fact-sheet-the-biden-harris-administrations-national-security-strategy/.

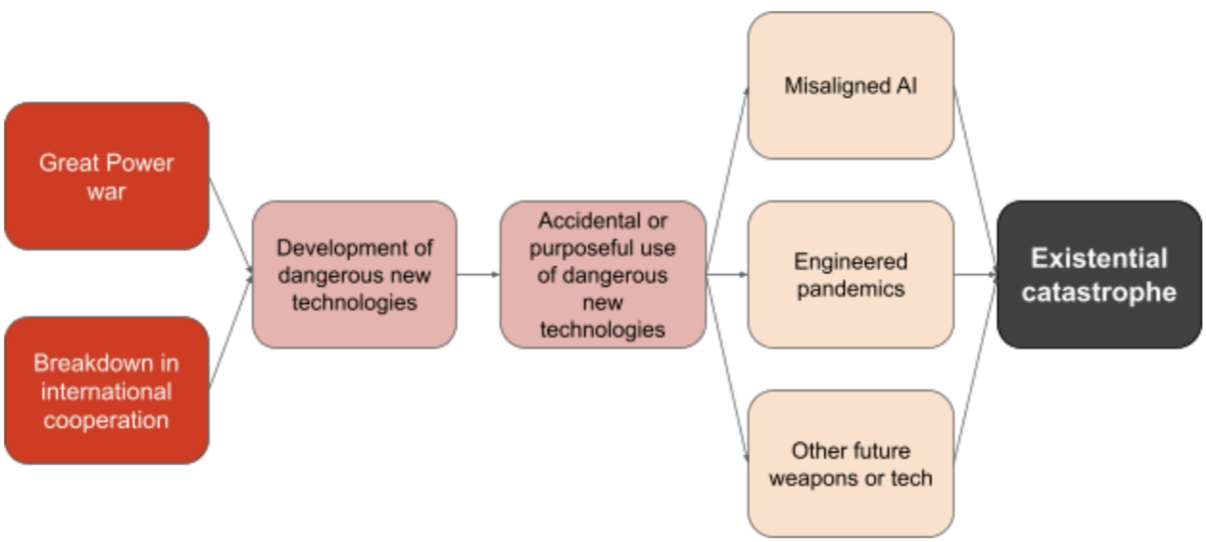
- **Military Applications and Procurement**: “As emerging technologies transform warfare and pose novel threats to the United States and our allies and partners, we are investing in a range of advanced technologies including applications in the cyber and space domains, missile defeat capabilities, trusted artificial intelligence, and quantum systems, while deploying new capabilities to the battlefield in a timely manner” (21)

Technology Competition Risk Pathways

In this report we show that more intense technology competition generally raises the level of existential risk humanity faces. This subsection describes a *risk pathways* model of how this occurs. Historical examples illustrate each risk pathway.

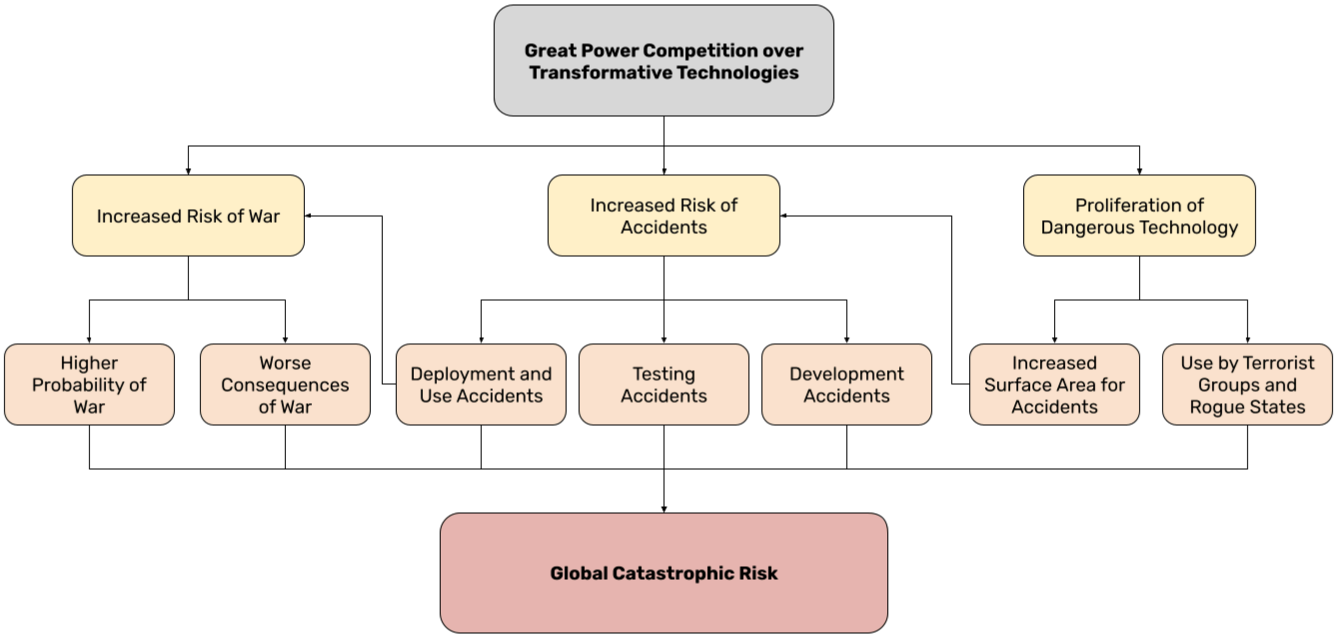
In Founders Pledge’s report on *Great Power Conflict*, Clare identified “great power war” and “breakdown in international cooperation” as two structural factors that may affect the development and use of “dangerous new technologies” such as transformative artificial intelligence, biotechnology, and other high-risk technologies:

Figure 1: Clare’s Outline of High-Risk Technology Competition



Source: Founders Pledge, *Great Power Conflict*, 78.

For analytical clarity, we can reframe and further disaggregate the potential pathways to global catastrophe:



This model shows seven distinct pathways by which intense technology competition could increase existential risk. These pathways fall into three categories.

First, technology competition may heighten international tension and raise the risk of war. It could make it more likely that a war in which dangerous technologies are used breaks out in the first place. Or it could increase the war-making capacity of the great powers, making a war more destructive.

Second, even if war doesn't break out, technology competition may increase the risk of military accidents which could themselves have devastating effects. These accidents could occur while technologies are being developed, while they are being tested, or while they are being deployed.

Third, technology competition may raise existential risk by driving proliferation of dangerous technology. It could drive more countries to try to obtain dangerous technologies, increasing the risk of accidents. As technologies proliferate, terrorists or other rogue groups may also find it easier to obtain them.

In the next section, we use this model to assess the "importance" of this problem in terms of how we expect technology competition to affect existential risk in coming decades.

Importance

The previous section outlined how great power competition over transformative technologies can affect the level of global catastrophic risk. This section assesses the importance of these dynamics

by estimating, roughly, how a given increase in the intensity of international technology competition affects the total existential risk faced by humanity.

This is obviously a difficult task. We try to make progress by:

- Studying the character, intensity, and focus of contemporary great power competition over transformative technologies
- Disaggregating "risk" into the risk of war, the risk of accidents, and the risk of proliferation, and studying each in turn
- Using historical case studies to better understand the complexity of this competition

In sum, we find that contemporary great power competition is focused on several transformative technologies, and that this competition appears likely to increase the potential severity of war, to increase baseline accident risk, and to increase the risk of dangerous technologies proliferating and falling into the wrong hands. We remain uncertain about the overall effect on the probability of war. As risk is a function of probability and consequence, the increased risk of great power war is therefore largely driven by the increased potential severity of war, enabled by transformative technologies.

Contemporary technology competition

One reason to take the risks of technology competition seriously is that we are already seeing evidence of competition over transformative technologies — or aspirations for technological superiority — in the world today. This section only briefly discusses the current state of competition to provide an overview.

First, the United States has explicitly declared technological superiority as a goal in various public strategic documents, as discussed above in the discussion of the National Security Strategy, and in the general U.S. strategic concept of "offsets." Similarly, the National Security Commission on AI, though not official policy, framed its recommendations in terms of technology competition, stating in its Final Report: "For the first time since World War II, America's technological predominance — the backbone of its economic and military power — is under threat. China possesses the might, talent, and ambition to surpass the United States as the world's leader in AI in the next decade if current trends do not change."[15] Technological "gaps" also feature heavily in discussions of the PRC's "pacing challenge" in U.S. national security circles. A 2019 Center for a New American Security (CNAS) report co-authored by Robert O. Work (often called the "father" of the Third Offset strategy), summarizes a point of view that appears to have become common in policy circles in recent years:

> The Soviets were never able to match, much less overcome, America's technological superiority. The same may not be true for China — certainly not for lack of trying. Indeed,

[15] National Security Commission on Artificial Intelligence, *Final Report* (Executive Summary), https://www.nscai.gov/wp-content/uploads/2021/03/Final_Report_Executive_Summary.pdf.

> China is keenly focused on blunting the U.S. military's technological superiority, even as it strives to achieve technological parity, and eventually technological dominance.[16]

There are many other examples of such language across government, like the director of the Defense Innovation Unit's statement that "We need technological advantage to prevail in this strategic competition with China [...] we're not going fast enough."[17] Notably, the Biden Administration's 2023 Executive Order on AI did not use this competitive framing, does not mention China, and the accompanying Fact Sheet emphasizes that "America already leads in AI innovation."[18]

Second, both Chinese and Russian leaders have made similar statements. Russian President Vladimir Putin famously said in 2017 that "Whoever becomes the leader in this sphere [AI] will become the ruler of the world."[19] Xi Jingping has highlighted China's need to "ensure that our country marches in the front ranks where it comes to theoretical research in this important area of AI, and occupies the high ground in critical and AI core technologies."[20] Some have argued that China has been pursuing a long-term "offset" strategy to match America's perceived superiority, and that this aim can be seen via a holistic analysis of Chinese actions and statements over the last decades, but it is important to note that "Chinese strategists do not explicitly describe their aims in this manner."[21]

Third, there is evidence of existing technological competition in a variety of areas, including but not limited to:

- Ongoing U.S. restrictions on China's access to high-end chips, the 2022 CHIPS Act, and related work on semiconductor manufacturing
- The DoD Replicator initiative to develop and field thousands of new AI-enabled military systems, framed explicitly as a strategy to counter China
- Private sector advocacy for greater technology competition, such as SCSP's Offset X strategy and related work
- Chinese testing of an apparent fractional orbital bombardment system (FOBS)
- Competition for high-end technology talent
    - China's Thousand Talents Program and High-End Foreign Expert Recruitment Program

---

[16] Robert O Work and Greg Grant, "Beating the Americans at Their Own Game," *CNAS*, 2019, https://www.cnas.org/publications/reports/beating-the-americans-at-their-own-game.
[17] "Tech Advantage Critical to Prevail in Strategic Competition With China, DOD Official Says," U.S. Department of Defense, accessed December 5, 2023, https://www.defense.gov/News/News-Stories/Article/Article/2835616/tech-advantage-critical-to-prevail-in-strategic-competition-with-china-dod-offi/.
[18] The White House, "FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence," The White House, October 30, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.
[19] "Whoever Leads in Artificial Intelligence Will Rule the World, Says Vladimir Putin," Fortune, accessed December 5, 2023, https://fortune.com/2017/09/04/ai-artificial-intelligence-putin-rule-world/.
[20] "Understanding China's AI Strategy," accessed December 5, 2023, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.
[21] Robert O Work and Greg Grant, "Beating the Americans at Their Own Game," CNAS, 2019, https://www.cnas.org/publications/reports/beating-the-americans-at-their-own-game.

    - Section 5 ("Promoting Innovation and Competition") in 2023 U.S. Executive Order on AI
- Executive Order on Advancing Biotechnology and Biomanufacturing Innovation
    - See by "Senior Administration Official" in Background Press Call on President Biden's Executive Order to Launch a National Biotechnology and Biomanufacturing Initiative that "Other countries, including and especially China, are aggressively investing in this sector, which poses risk to U.S. leadership and competitiveness unless we take the kinds of actions that we are with this executive order"

Technology competition will likely be intense in coming decades

The great powers are already competing on various potentially transformative technologies through export controls, immigration policy, defense R&D, bureaucratic reorganization, programs of record, and in other ways. However, as mentioned in the previous section, technology competition can vary in intensity. The risk it poses varies similarly.

To assess its importance going forward, therefore, we need to consider how the intensity of competition over transformative technologies may change over the coming decades. Is it likely to slowly escalate, to rapidly escalate, to remain roughly at the current level, or to decrease in intensity?

This is, of course, a tough question. While forecasting science has advanced in recent years, even experienced forecasters often remain highly uncertain about the future of geopolitical trends. We have very little data on how accurate forecasters are on timescales of about 10 years or more. And forecasting questions have specific resolution criteria that make it hard to apply them to big-picture questions like "the future of technology competition," although there have been some attempts to harness the power of forecasting for such questions (e.g. Center for Security and Emerging Technology's work on Future Indices and Foretell).

We think these challenges reflect hard-to-reduce uncertainties about geopolitical trends. Dramatic geopolitical shifts can occur suddenly. A range of future scenarios are possible, from lower tensions and renewed cooperation to serious and violent escalation.

That said, we do not think we should consider all these scenarios *equally* possible. Instead, it seems most likely that technology competition will intensify in the coming decades. This seems likely for a few reasons:

- **Political signals:** political actors in all great powers tend to express the notion that they are in competition
- **Conflict spirals:** competition among the great powers has escalated recently, and escalatory actions tend to provoke escalatory reactions

- **Commitment problems:** the relative rise of China as a share of world GDP, and potential future rise of India, present commitment problems; as some nations' relative power grows their leaders may wish to renegotiate agreements on more favorable terms
- **Intractable issues:** some of the issues raising tension among great powers seem unlikely to be resolved soon (e.g. sovereignty of Taiwan)
- **Security dilemmas:** a security dilemma refers to a situation where state A's defensive actions to increase its security create feelings of insecurity in a rival state B, which in turn reacts by preparing for the worst interpretation of state A's actions, ultimately undermining everyone's security and sometimes sparking arms races[22]
- **Heightened stakes:** more speculatively, the transformative nature of some emerging technologies could heighten competitive dynamics

The potential for conflict spirals and existence of intractable issues, in particular, mean that we cannot rule out the possibility that these tensions could lead to direct conflict. Recent estimates have put the chance of some kind of direct conflict between great powers before 2050 at one in three.[23] Most wars remain relatively small, but there's some chance that it could escalate enormously and come to equal or surpass the devastation of World War II.

Some relevant crowdsourced probabilistic forecasts inform this picture of heightened tensions:

Forecast Question	**Probability**[24]	**Notes**	**Platform**
Will there be a U.S.-China war before 2035?	15% (annualized 1.35%; implied 30.7% by 2050)	Resolves on 1,000 battle-related deaths in a calendar year. 288 forecasters with 528 predictions.	Metaculus
Will there be active warfare between the	7%	Resolves on news reports of kinetic	Metaculus

[22] The classic statement of the security dilemma is John H. Herz, "Idealist Internationalism and the Security Dilemma," *World Politics* 2, no. 2 (1950): 157–80, https://doi.org/10.2307/2009187. See also Charles L. Glaser, "The Security Dilemma Revisited," *World Politics* 50, no. 1 (1997): 171–201. Ken Alibek, a Soviet Defector and high-level scientist in Biopreparat, later reflected that U.S. defensive activities made them question Nixon's sincerity in renouncing biological weapons:

> "[W]e also noted that a small army medical unit had begun work at Fort Detrick. This unit, known as the United States Army Medical Research Institute of Infectious Diseases and ostensibly dedicated to biological defense, seemed to expand in importance and strength each year. Former bioweaponeers like Bill Patrick had gone to work for it. Even if our intelligence activities couldn't come up with concrete evidence of offensive work, there could be no doubt that such work continued."

Alibek, *Biohazard: The Chilling True Story of the Largest Covert Biological Weapons Program in the World-Told from Inside by the Man Who Ran It*, 235.

[23] See "How Likely is a War" in Stephen Clare, "Great Power War: Problem Profile" (80,000 Hours, 2023), https://80000hours.org/problem-profiles/great-power-conflict/.

[24] As of December 2023.

United States and China before 2027?	(annualized 1.8%; implied 38.8% by 2050)	warfare. 264 forecasters with 602 predictions.	
Will the US enact export controls for some generative AI software before 2026?	75%	Resolution not restricted to China. 64 forecasters with 106 predictions.	Metaculus
Will any TSMC fab in Taiwan be shut down for at least one day because of a non-scheduled emergency by 31 March 2024?	1.57%	Can include natural disasters. 71 forecasters with 389 predictions.	INFER-pub
Will there be a Frontier AI lab in China before 2026?	80%	Frontier lab defined as "one who has trained models within one order of magnitude of the largest known model." 93 forecasters with 183 predictions.	Metaculus

All that said, there are also a few considerations that give us pause, suggesting that it's also possible that the great powers find ways to lower tensions and renew cooperation in the future:

- **Precedent for détente:** the Cold War ended without the US and the USSR coming into direct conflict; it seems like leadership transitions in autocratic nations provide particular opportunities for "resets" and easing of hostilities
- **Long-term trend towards international cooperation and peace:** recent decades have seen international institutions proliferate; these seem likely to have some pacifying effect by providing diplomatic channels and interdependencies, though it's not clear how strong that effect may be
- **Heightened risks:** as a balance to the heightened stakes point, the global risks posed by transformative technologies could drive countries to cooperate because the expected costs of competition are so large

There are strong considerations either way, justifying uncertainty. But on the whole the trend seems to be towards intensifying geopolitical competition.

How will competition affect existential risk?

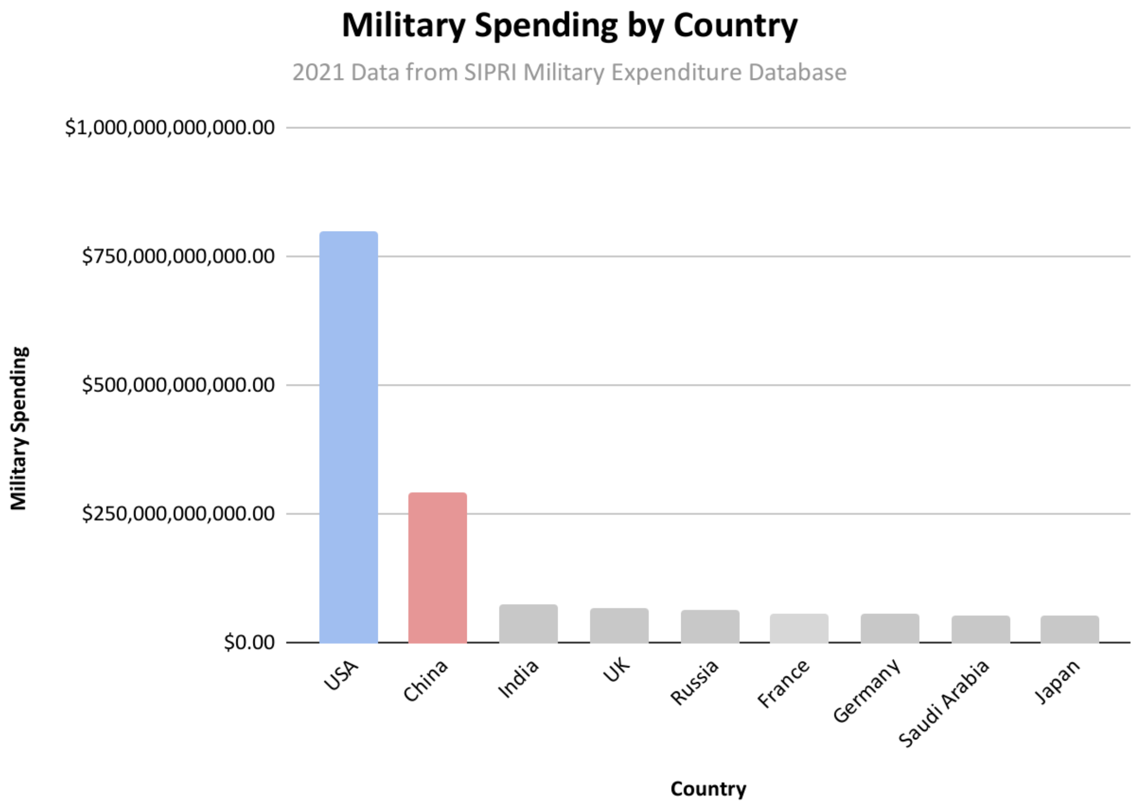
While it is very difficult to predict how geopolitical events will unfold, we have seen that there are some reasons to expect some intensification of technology competition in the coming decades — as many policymakers clearly do. The next critical question to consider is how this intensification will affect the level of existential risk we face.

We first discuss an important underlying factor motivating this report: the resource leverage of great power governments. We then return to the risk pathways model introduced in the previous section to discuss how technology competition could affect the probability of various global catastrophes.

Resource leverage

Governments of major military powers control many resources – including large R&D budgets and talent stocks– that they can devote to the development of transformative technologies. Under intense technology competition, these resources could accelerate the development of such technologies, or perhaps lead more dangerous technologies to be invented than otherwise would have been.

The United States and China spend hundreds of billions of dollars on their militaries every year, dwarfing the spending of the next closest states:



Source: Author's diagram using SIPRI Military Expenditure Database.

They devote a significant fraction of these budgets to developing new technologies. For example, the US 2023 budget request included $204.9 billion for research and development.[25] This is about the same as *total* US venture capital funding, which, according to industry research firm Crunchbase, in 2022 totalled about $215 billion.[26] This financial heft could give governments significant influence over technology research priorities. Moreover, the level of military spending, as well as its allocation, varies over time.[27] Should technology competition intensify, total spending could increase and more of it could be devoted to R&D. Finally, public R&D spending can stimulate additional private funding: one recent analysis found that, on average, "a 10% increase in government-financed R&D generates a 5% to 6% additional increase in privately funded R&D."[28]

---

[25] Congressional Research Service, "Federal Research and Development (R&D) Funding: FY2023" (Congressional Research Service, 2022), https://crsreports.congress.gov/product/pdf/R/R47161.

[26] Albeit this is down almost 40% from its peak of nearly $350 billion the previous year. Joanna Glasner, "AI's Share Of US Startup Funding Doubled In 2023," Crunchbase News, August 29, 2023, https://news.crunchbase.com/ai-robotics/us-startup-funding-doubled-openai-anthropic-2023/.

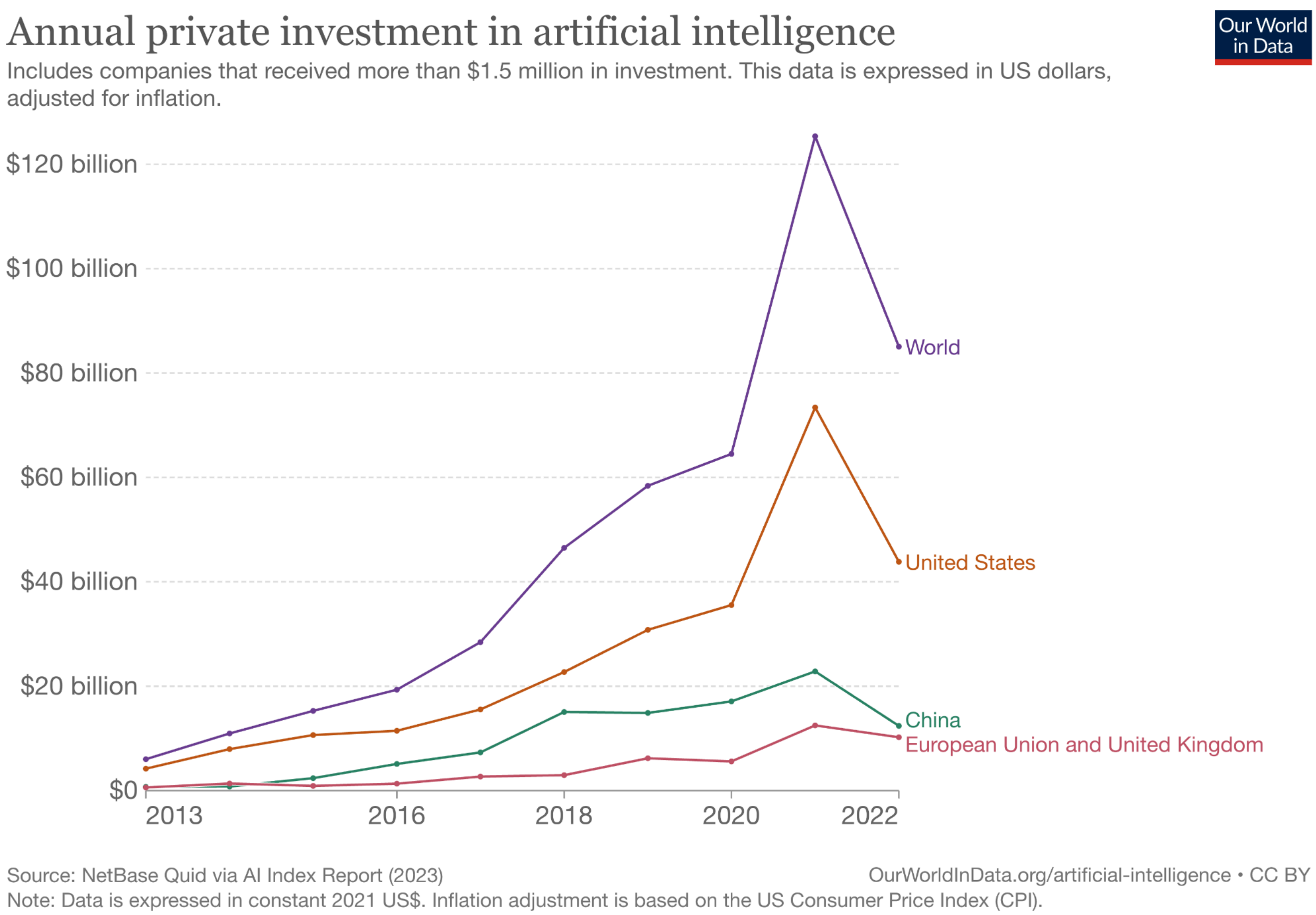
[27] "Military Expenditure as a Share of GDP," Our World in Data, accessed January 12, 2024, https://ourworldindata.org/grapher/military-expenditure-share-gdp?tab=chart&country=USA~RUS~CHN~IND.

[28] Enrico Moretti, Claudia Steinwender, and John Van Reenen, "The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers," Working Paper, Working Paper Series (National Bureau of Economic Research, November 2019), https://doi.org/10.3386/w26483.

Governments can also leverage resources through regulatory actions. Consider the example of machine learning and other compute-intensive technologies. American subsidies and export/import controls helped shape the global semiconductor supply chain during the Cold War and 1990s. More recently, the US government has tried to restrict Chinese actors' access to advanced computer chips by banning the sale of semiconductor manufacturing products which use American technologies.[29]

The power to shape the behavior of private actors gives great power governments another powerful lever. And again, the US and China lead the world in venture capital investments in important sectors like artificial intelligence:

Venture Capital Investments in AI by Country (OECD.AI)

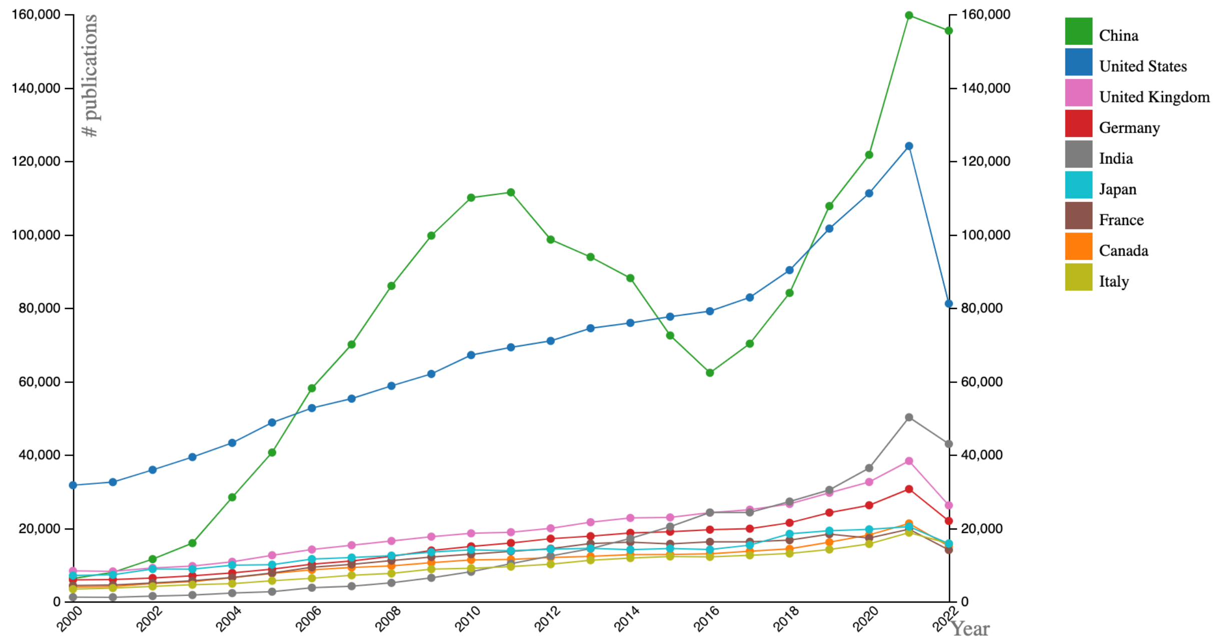


Source: Our World in Data, https://ourworldindata.org/grapher/private-investment-in-artificial-intelligence

[29] For an overview of U.S. government involvement in the semiconductor supply chain, see *Chip War.* On the specific regulations of late 2022, see Bureau of Industry and Security, "Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)," 2022, https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file.

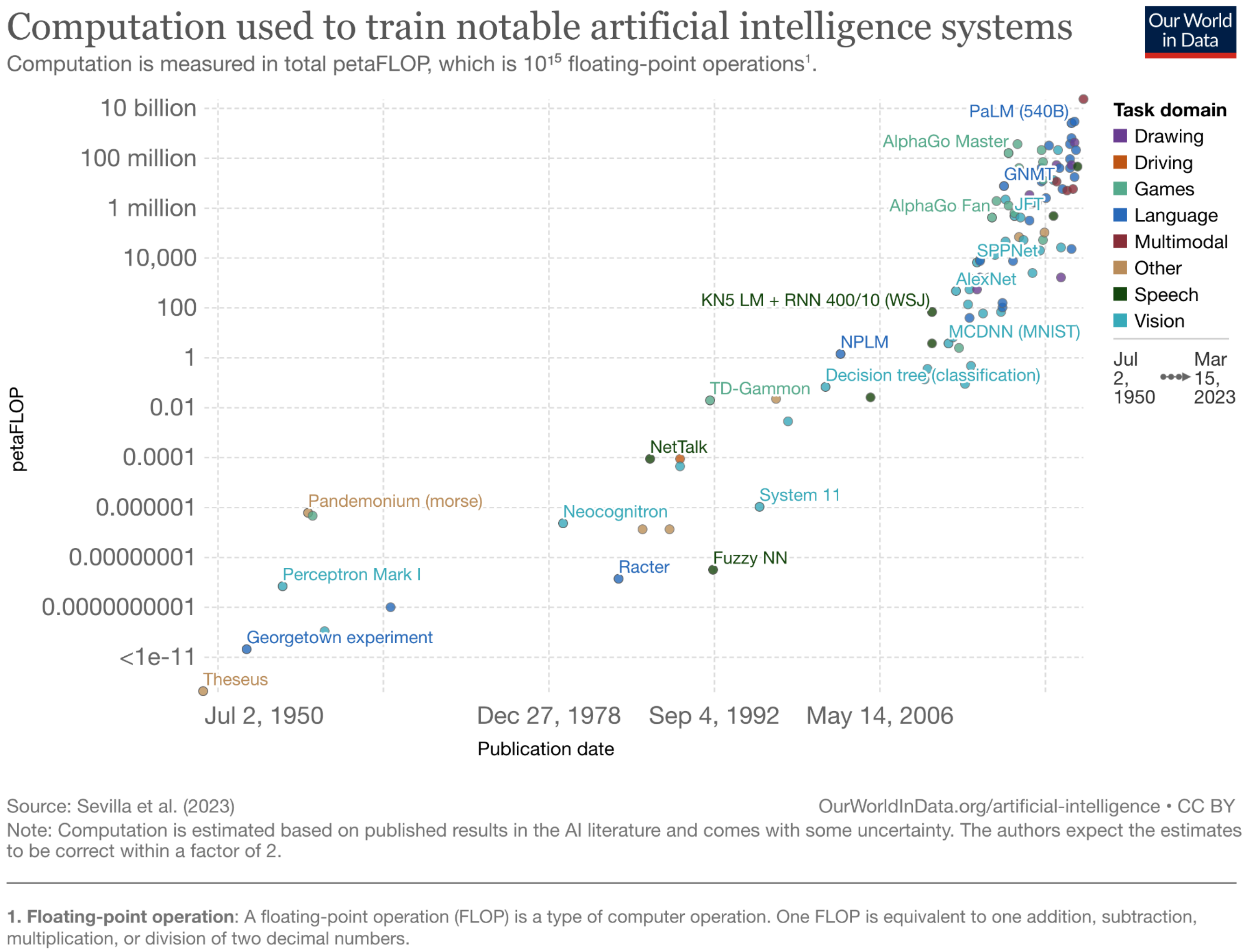
AI publications display a similar pattern:

AI Publications by Country (OECD.AI)



Source: OECD.AI (2023), visualisations powered by JSI using data from OpenAlex., accessed on 20/3/2023, www.oecd.ai

This matters because the development of high-risk technologies can require immense resources. This is generally appreciated for nuclear technology, where a decades-old non-proliferation regime and the difficulty of acquiring weapons-grade fissile material is well-known. In the current approach to artificial intelligence, computational resources ("compute") may be similarly important:



Crucially, **competition on frontier AI models has largely been driven by the private sector so far** (i.e. competition between AI labs and their corporate backers). We believe that the risk of government-driven AI competition remains an under-examined driver of far more intense competition. If great powers truly throw their resources and might behind developing more cutting edge models, what appears as intense competition now may pale in comparison to major government efforts. Therefore, **philanthropists concerned about AI risk should take more seriously the possibility of great-power competition as an accelerant.**

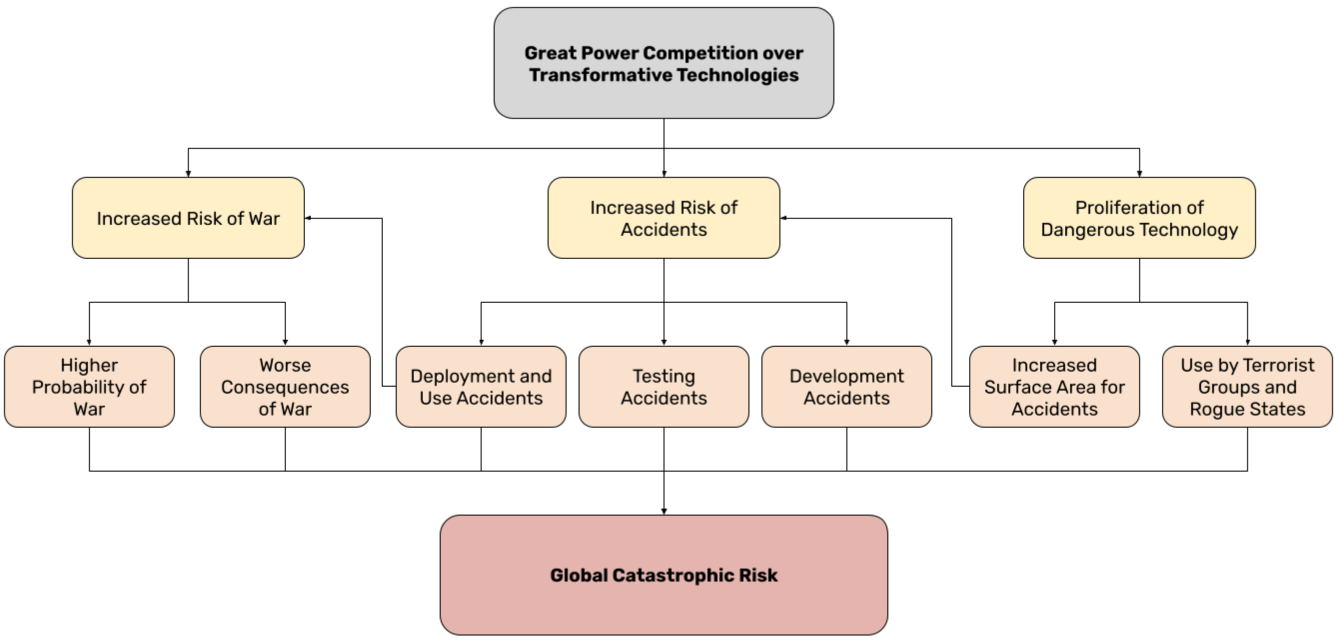
To summarize this section:

- Heightened international tensions could lead great powers to increase military spending
- Since technological superiority is a key goal of modern militaries, those governments may also devote a larger fraction of their military budgets to technology R&D

- Military R&D budgets already rival private investment in new technologies as measured by, for example, total venture capital funding
- This means that heightened international tension could substantially increase the total amount of funding devoted to developing new technologies
- Governments can also influence private funding through regulations, such as subsidies, tax laws, and controls on imports and exports
- In other words, great power governments, especially in the US and China, have enormous resource leverage

Risk pathways

To assess this question in more detail we apply the "technology competition risk pathways" model introduced in the previous section. We assess each pathway in term, considering (and quantifying, where possible) the effect we expect intensifying competition to have on existential risk.



Our results are summarized in the following table:

Threat Pathway	**Net Effect of Technology Competition**
<u>Probability of War</u>: How likely are major powers to engage in military conflict?	**Effect Unclear or Small** — Although there are plausible pathways for qualitative arms racing to increase the probability of war, new technologies may have powerful deterrent effects or other stabilizing effects, and there are empirical uncertainties and deep

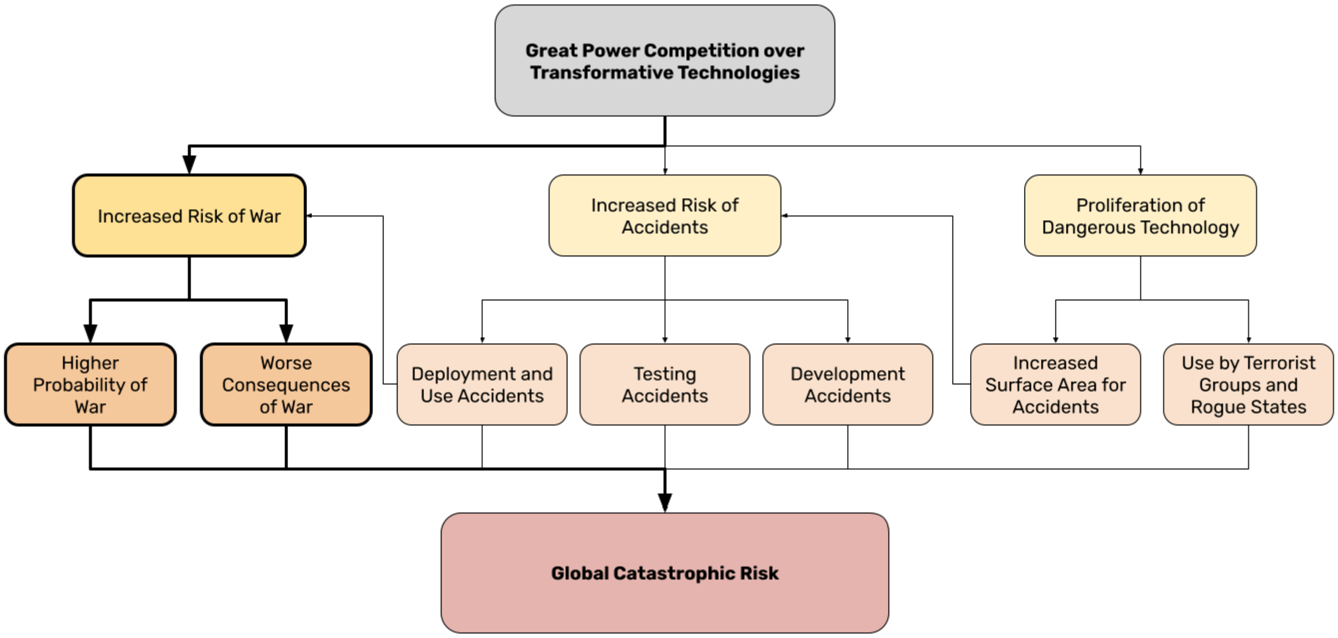
	methodological challenges with studying this question. We believe that a small effect may exist.
Destructiveness of War: Does technology competition make wars more severe?	**Increased Risk** — Although recent military innovation has favored precision weaponry, transformative technology competition may unleash new types of weapons of mass destruction, whose effects we expect to outweigh other trends.
Accidents — Does competition increase the risk of accidents in technology development, testing, or deployment?	**Increased Risk** — Militaries have strong incentives to avoid using uncontrollable weapons, but race-to-the-bottom dynamics and historical examples suggest that safety trades off against speed in the development of transformative technologies.
Proliferation — What happens when transformative technologies fall into the wrong hands?	**Increased Risk** — As more countries pursue transformative technologies, leaks and theft become more likely; as with accident risk, information security and nonproliferation may take a backseat to "winning the race." Once a dangerous technology is developed, there is a risk of it being obtained by terrorist groups or rogue states who could use it for mass-casualty attacks.

Will strategic competition over transformative tech raise risk from war?

We first examine whether technology competition will raise the risk from war, defined as the product of the **probability of a war** and its **consequences**. We conclude that **the effect on the risk of war is ultimately unclear**. Although technology competition can certainly raise international tensions and accelerate the development of extremely dangerous weapons, some weapons may have deterrent effects. It is very difficult to estimate the relative strength of these counterbalancing effects, leaving us uncertain about the net impact.

There is a large academic literature on arms races and war. However, most of it focuses on *quantitative* arms races rather than *qualitative* arms races like competition over transformative technologies. Moreover, it is largely theoretical and relies on game theory models in which states are portrayed as unitary actors. Little empirical work exists (this is unsurprising, given the methodological challenges of running empirical studies on war risks).

While these challenges leave us with substantial uncertainty about the effect of technology competition on total war risk, we do believe that technological competition is likely to increase the risk of military accidents. Such accidents can lead to inadvertent escalation. Although this pathway leads to disaster through the effects of a war, we discuss it in the section on accident risks.



Probability of war

In theory, great power competition over transformative technologies could make a major war more likely. Since a modern great power war could become an unprecedented catastrophe,[30] this would make high-risk technology competition a risk factor for global catastrophe.

To examine this claim, we can look at one subset of strategic competition over transformative technologies: "qualitative arms races." Unfortunately there is no expert consensus on whether qualitative arms races make wars more likely.[31] Endogeneity makes this a difficult question to study. Say we observe that it often does end in war: should we conclude that competition makes wars more likely, or that the same factors driving countries to compete also lead them to war?[32]

Lacking strong empirical evidence, we can look to theoretical work linking technology competition to war. Researchers have produced a large array of models that try to explain why and when wars occur. Again, though, these models prove difficult to verify empirically. As a result there is little

---

[30] Stephen Clare, "Great Power Conflict" (Founders Pledge, 2021), https://dkqj4hmn5mktp.cloudfront.net/Great_Power_Conflict_report_Founders_Pledge_e4124df2ac.pdf

[31] For an overview of this literature, see Richard J. Stoll, "To Arms, To Arms: What Do We Know About Arms Races?," in *Oxford Research Encyclopedia of Politics*, September 26, 2017, https://doi.org/10.1093/acrefore/9780190228637.013.350.

[32] "[T]he security environments that require states to engage in military competition are also more likely to generate insecurity and war than are those that allow states to pursue more cooperative arming policies" and "Finding a correlation between arms races and the probability of war may say little about the impact of arms races, because the correlation could simply reflect the causal impact of dangerous security environments" Glaser, "When Are Arms Races Dangerous? Rational versus Suboptimal Arming," *International Security* 28, no. 4 (2004): 47.

consensus over which are most useful. Here we focus on two which seem particularly relevant: the **spiral model** and the **deterrence model**. These examine the effects of new technologies on power balances and the ability of defenders to respond effectively to attacks, respectively. These models are typically used to analyze *quantitative* arms races, but they may also apply to qualitative races.[33]

In the spiral model, efforts by one state to develop transformative technologies may spur increased efforts from a rival, leading to a spiral of ever-increasing investments. Spirals are particularly dangerous when it is difficult to differentiate defensive technologies from offensive ones. Importantly, spirals can also lead to a deterioration of trust between rivals as they perceive each other as taking aggressive actions. Technology competition can push states to take increasingly aggressive actions, like instituting trade policies designed to undermine a rival's technological, and thus economic, progress. This may spark escalatory spirals that bring the rivals closer to war.

The spiral model suggests that heightened competition is likely to increase the risk of war. For instance, in the coming decades both the US and China may invest heavily in AI-enabled aerial and maritime weapon systems as force postures become more aggressive and the Indo-Pacific begins swarming with thousands of uncrewed systems. In these conditions, a single accident with one autonomous ship or swarm may spark a rapid series of escalations, potentially leading to war.[34]

The deterrence model focuses on the possibility that technology competition could accelerate the development of destabilizing technologies. The invention of a revolutionary military technology could cause an actual or perceived change in the balance of power. Such power transitions may be especially dangerous times (though the empirical evidence is mixed, as it is for all theories of war). If leaders in the newly-powerful state calculate that the odds are now in their favor, they could be incentivized to launch an attack to achieve a certain objective.

For instance, in the same scenario described above, China may perceive that its integration of artificial intelligence into its military forces, combined with its growing nuclear arsenal, gives it a temporary window of opportunity in the Taiwan Strait, and launches an invasion of Taiwan.

The deterrence model suggests technology competition will only have a small effect on the likelihood of war. Heightened competition could generally accelerate the development of new military technologies, and these technologies may be destabilizing. But they could also make the world more stable. For example, they could differentially benefit defenders rather than attackers, making it harder to fight successful wars.[35]

---

[33] We discuss below whether studies of quantitative arms racing have much explanatory power for the competition discussed in this report.

[34] Scharre, "A Million Mistakes a Second," Foreign Policy, 2018, https://foreignpolicy.com/2018/09/12/a-million-mistakes-a-second-future-of-war/.

[35] Ben Garfinkel and Allan Dafoe, "How Does the Offense-Defense Balance Scale?," 2019, https://web.archive.org/web/20200305192814/https://www.fhi.ox.ac.uk/wp-content/uploads/How-Does-the-Offense-Defense-Balance-Scale-5-21-19.pdf.

In the biotechnology field, for instance, competition could drive the development of strong defensive technologies — germicidal ultraviolet light, broad-spectrum medical countermeasures, pandemic-proof PPE, and other pathogen-agnostic defensive technology — that make it harder to carry out successful biological attacks. (*We emphasize that we do not think this is an accurate description of what would happen in great power competition over weaponized synthetic biology, which could be extremely dangerous, as outlined in* Global Catastrophic Biological Risks: A Guide for Philanthropists.) Similarly, strategic competition over transformative technologies may sometimes provide a more peaceful outlet for great power tension, satisfying the demands of bellicose constituents (public or bureaucratic), who would otherwise clamor for more belligerent actions.

Given the weak evidence base and multitude of considerations, we remain uncertain about the effects of technology competition on the probability of war. It seems more likely than not that intense competition makes war somewhat more likely in expectation. But given the unpredictable effects of new technologies, and the myriad of factors that influence the decision to go to war, it's probably not a very strong effect.

As we've argued above, technology competition in the coming decades seems likely to be relatively intense. And previous estimates have put the current probability of a great power war before 2050 at about 30%.[36] By how much could this be lowered if, instead, we experience a renewal of cooperation between great powers and, at least in the domain of emerging technologies, international agreements to limit shared risks? It is impossible to say for sure. And we want to emphasize that this report concerns *technology* competition specifically; we could still see hostilities in other domains, such as conventional military clashes over disputed territories or trade disagreements.

But it does seem plausible that some fraction of the overall risk could be lowered by reducing technology competition. Agreeing to, for example, shared safeguards in AI development and prohibitions on research into new WMDs like destabilizing nuclear weapons technology or bioweapons seems likely to close off a few pathways to war. We would probably be somewhat less likely to enter a conflict spiral or see political leaders willing to undertake risky gambes. If we conservatively forecast that this could make a war 5% less likely and apply it to the previous estimate of a 30% chance of war before 2050, this would lower the overall chance of conflict before 2050 by between one and two percentage points.

---

[36] "my forecast is that there's about a 30% chance we see a conflict that technically qualifies as a war (i.e. involves at least 1,000 battle deaths in a year) involving at least one of the United States, China, Russia, and India on each side before 2050. 30% is not my forecast of the chance that we see "World War III", widespread use of nuclear weapons, or some similar kind of global catastrophe. I think these events are much less likely." Stephen Clare, "Great Power War: Problem Profile" (80,000 Hours, 2023), https://80000hours.org/problem-profiles/great-power-conflict/.

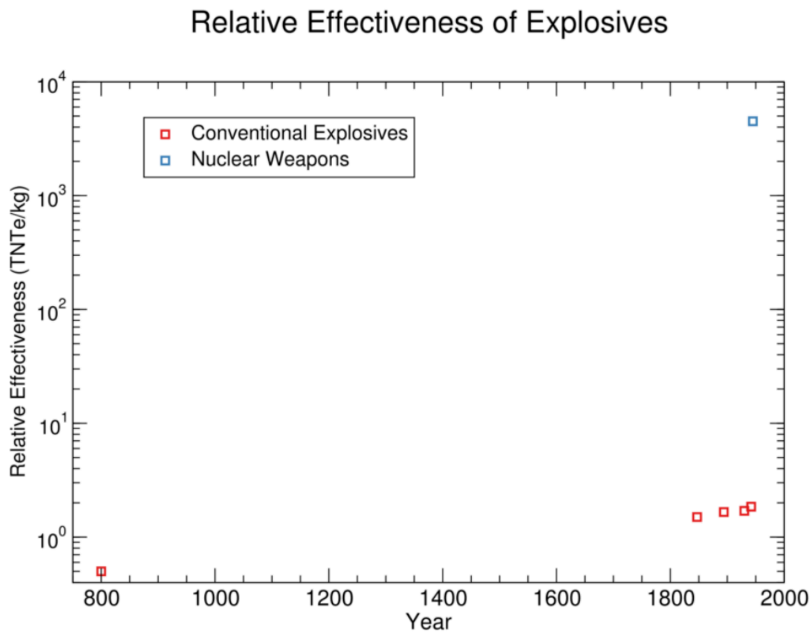
Case Study: Cold War Nuclear Arms Racing

One example that outlines the ambiguity of arms races on the probability of war is the case of Cold War nuclear arms racing between the United States and the Soviet Union. Though partly quantitative, the Cold War buildup also had elements of qualitative arms races, including competition over delivery systems, bomb design, and missile defense systems. Some of this competition found outlets in purportedly civilian competition, like the Space Race. Several attempted empirical studies seem to suggest, for example, that arms races did *not* increase the probability of conflict during the Cold War.[37] Nonetheless, some competition over potentially transformative technologies *did* clearly cause instability. For example, Reagan's pursuit of the Strategic Defense Initiative appears to have set off a period of heightened tensions, where "Soviet sources [...] warned of 'growing paranoia among Soviet officials," whom the source described as "literally obsessed by fear of war.'"[38] This paranoia appears to have contributed to the Able Archer "war scare" of 1983 (although there is scholarly debate about the existence of this "war scare itself").[39]

Severity of war

Probability is just one factor influencing the risk we face from war; we also have to consider its severity. And technology competition could make war more destructive.[40] The invention of new weapons can cause large discontinuities in the destructiveness of war. Nuclear weapons, for example, were a step-change in humanity's war-making capacity:

[37] As Stoll summarizes the studies, for example, Senese and Vasquez (2008) conclude that "For the Cold War period, mutual arms buildups do not increase the chances of war" and Sample (2012) that "a mutual arms buildup in the pre-Cold War era is likely to lead to an escalation of a militarized interstate dispute (or any dispute occurring over the following five years) to war. But this relationship does not occur during the Cold War." (Richard J. Stoll, "To Arms, To Arms: What Do We Know About Arms Races?," in Oxford Research Encyclopedia of Politics, September 26, 2017, https://doi.org/10.1093/acrefore/9780190228637.013.350.)
[38] National Security Archive, "The 1983 War Scare: 'The Last Paroxysm' of the Cold War Part I," accessed September 27, 2023, https://nsarchive2.gwu.edu/NSAEBB/NSAEBB426/.
[39] For a discussion of this, see Simon Miles, "The War Scare That Wasn't: Able Archer 83 and the Myths of the Second Cold War," Journal of Cold War Studies 22, no. 3 (August 1, 2020): 86–118, https://doi.org/10.1162/jcws_a_00952.
[40] In this case, we use "destructiveness" to refer not only to battle deaths, but to broader damages, including deaths, suffering, grief, economic damage, infrastructure damage, knowledge loss, and more.



Source: AI Impacts, "Effect of nuclear weapons on historic trends in explosives"

There are a few reasons to think that technology competition could increase the severity of a great power war.

The first is that the *long-term* trend of military innovation has been towards successively more destructive weapons. Machine guns are more destructive than bows, and nuclear bombers are much more destructive than catapults. Humanity's war-making capacity has greatly increased over time.[41] This is partly because economic growth has allowed states to build larger arsenals (and armies and war-supporting industry), but the much stronger effect appears to be technological change and the invention of more powerful weapons. Although militaries don't pursue lethality *per se*, lethality is positively correlated with their other goals.

The second reason is that experts believe there is a good chance that various powerful technologies will be invented this century.[42] These technologies are likely to have important military applications. As discussed in our other research reports, these may include:

- Research into bioweapons leading to the creation, intentional or not, of engineered pathogens with the potential to kill large swathes of humanity

[41] See Stephen Clare, "Great Power Conflict" (Founders Pledge, 2021), https://dkqj4hmn5mktp.cloudfront.net/Great_Power_Conflict_report_Founders_Pledge_e4124df2ac.pdf.
[42] John Halstead, "Safeguarding the Future" (Founders Pledge), accessed January 12, 2024, https://www.founderspledge.com/research/existential-risk-executive-summary.

    - For discussion of counterarguments that states will see such weapons as tactically or strategically useless, see our report on Global Catastrophic Biological Risks and the Council on Strategic Risks report "Common Misconceptions about Biological Weapons"
- Research into military artificial intelligence systems, able to control large arsenals of autonomous weapons at lightning speed, which operators may not be able to perfectly control
    - See Founders Pledge's report on Autonomous Weapon Systems and Military AI
- Next-generation nuclear weapons or a significant expansion of existing nuclear arsenals
- Other weapons whose invention is currently hard to predict
    - E.g. new advances in directed energy weapons, information warfare, psychological warfare, and other advanced weapon systems

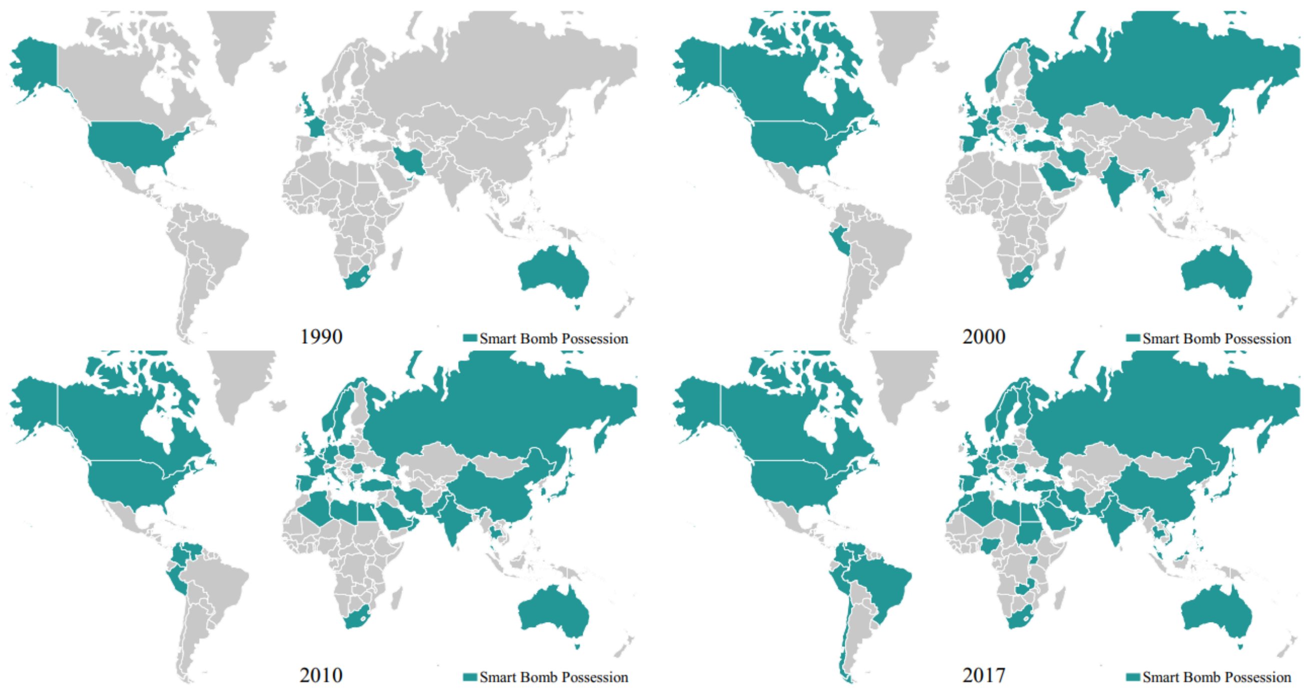
Strategic competition could accelerate the development of these weapon systems, which are all enabled by transformative technologies. More speculatively, it could cause some technologies to be invented when they otherwise would not have been. During the Cold War, for example, the Soviets operated a covert bioweapons program in part because they feared the Americans were doing the same. Had relations been better, the Soviets would probably have been more likely to disband their program along with the US in the 1970s. This would certainly have slowed bioweapons research and may have basically brought it to a halt.

The third is that intense technology competition may lead states to increase their military budgets generally, allowing them to build larger arsenals and fight larger wars. (Notably, however, technological competition is sometimes framed as a resource-saving measure that delivers "more bang for the buck", as the discourse around various technology-focused "offset strategies" in the United States illustrates.)

It is not the case, though, that new technologies *always* make war more destructive. They may also, for example, decrease the risk of collateral damage or help defend against attacks.[43] Militaries do not pursue lethality *per se*, but specific objectives like controlling territory and winning political disputes. Some of these objectives, like gaining the support of citizens in occupied territories, may be unrelated or even opposed to increased lethality.

[43] "It seems like states will also be more likely to develop defensive technologies if international tension is high. These defenses could be helpful in mitigating accidents as well as purposeful attacks. So if the probability that an accident is avoided outweighs the probability that an attack is successful, a more tense world could be a safer one." Stephen Clare, "Modelling Great Power Conflict as an Existential Risk Factor," Effective Altruism Forum, 2022, https://forum.effectivealtruism.org/posts/mBM4y2CjfYef4DGcd/modelling-great-power-conflict-as-an-existential-risk-factor.

Indeed, in recent decades much military R&D has focused on increasing the precision of weapons rather than their destructiveness.[44] This has led to the proliferation of so-called "smart bombs" or precision-guided munitions:



Source: Lauren Kahn and Michael Horowitz, "Who Gets Smart? Explaining How Precision Bombs Proliferate"

This dynamic makes the net effect of technology competition on war severity somewhat more difficult to predict. Nonetheless, we should not over-fit to this more recent trend. Indeed, increased precision need not entail decreased severity of war (or it may decrease the severity of regional interventions, but not of major wars between great powers). Indeed, the apparent benefits of increased precision may turn out to be a mirage if they lead decision-makers to overestimate the controllability of certain weapon types. For example, if advances in the life sciences enabled the development of more biological weapons targeted to specific genetic profiles (sometimes called "genetic warfare"), states may pursue contagious pathogens as assassination weapons (e.g. for

[44] For an overview, see Lauren Kahn and Michael C. Horowitz, "Who Gets Smart? Explaining How Precision Bombs Proliferate," *Journal of Conflict Resolution* 67, no. 1 (January 1, 2023): 3–37, https://doi.org/10.1177/00220027221111143.

targeted strikes against insurgent leaders), thereby weakening the taboo against biological warfare in general, including weapons of mass destruction.[45]

On the whole, the long-term trend and expert views on emerging transformative technologies lead us to think technology competition is likely to make wars somewhat more destructive (in expectation) in the coming decades. Again, less intense competition could practically look like more safeguards on the development of new military technologies and perhaps bans on particularly risky applications, like developing biotechnology techniques or integrating AI systems into nuclear command and control. Limiting the diffusion of such technologies seems more likely to limit the severity of future wars than increase it. But this effect seems very difficult to measure, and existing weapons alone could make a modern great power war extraordinarily destructive.[46]

Case Study: Megaton-Range Bombs

The case of competition over very large thermonuclear weapons with yields above 1 megaton (MT) of TNT helps to illustrate the difficulty of understanding the effects of military competition on the destructiveness of war. US-Soviet nuclear competition drove the initial development of the hydrogen weapons. Some American atomic scientists advised strongly against the development of such weapons, writing in the General Advisory Committee's 1949 report, "the existence of a weapon of this type whose power of destruction is essentially unlimited represents a threat to the future of the human race which is intolerable."[47] The opposite camp, whose arguments won out in the Truman administration, couched its arguments in terms of keeping up with Soviet technical capabilities: "Recent accomplishments by the Russians indicate that the production of a thermonuclear weapon is within their technical competence."[48] Ultimately, this competition led to the development of massive weapons, like the largest nuclear bomb ever developed, Tsar Bomba, which had a yield 3,800 times as large as the bomb dropped on Hiroshima.[49] This pursuit of ever-larger weapons did not, however, continue unabated. Despite much talk about competition with China and Russia, the United States has not pursued larger nuclear weapons in recent years, and the 2022 Nuclear Posture review outlines a plan for the retirement of the last megaton-range weapon in the US arsenal (there has been pushback from Congressional Republicans).[50]

---

[45] Yelena Biberman, "The Technologies and International Politics of Genetic Warfare," *Strategic Studies Quarterly,* https://www.airuniversity.af.edu/Portals/10/SSQ/documents/Volume-15_Issue-3/Biberman.pdf.
[46] See "War Could Be Devastating" in Stephen Clare, "Great Power War: Problem Profile" (80,000 Hours, 2023), https://80000hours.org/problem-profiles/great-power-conflict/#war-could-be-devastating,
[47] General Advisory Committee to the U.S. Atomic Energy Commission, Report of October 30, 1949.
[48] For the full reasoning, see Lewis Strauss's November 25, 1949, Memorandum to President Eisenhower.
[49] Amy Tikkanen, "Tsar Bomba," Encyclopedia Britannica, 6 Jan. 2023, https://www.britannica.com/topic/Tsar-Bomba. Accessed 19 April 2023.
[50] Bryant Harris, "Republicans Lay Battle Lines over Biden's Plan to Retire B83 Megaton Bomb," Defense News, May 19, 2022, https://www.defensenews.com/congress/budget/2022/05/19/republicans-lay-battle-lines-over-bidens-plan-to-retire-b83-megaton-bomb/.

Summary of effect on war risk

We have so far examined two potential effects of intensifying technology competition. First, whether it will make wars between great powers more likely to occur; and second, whether it will make any such wars more destructive should they occur.

Each of these questions proves difficult to answer with confidence. The future of technological development is deeply uncertain. And further, it is difficult to draw strong lessons from the history of technology competition and war. As a result, we caution against overconfident views about the likely effects.

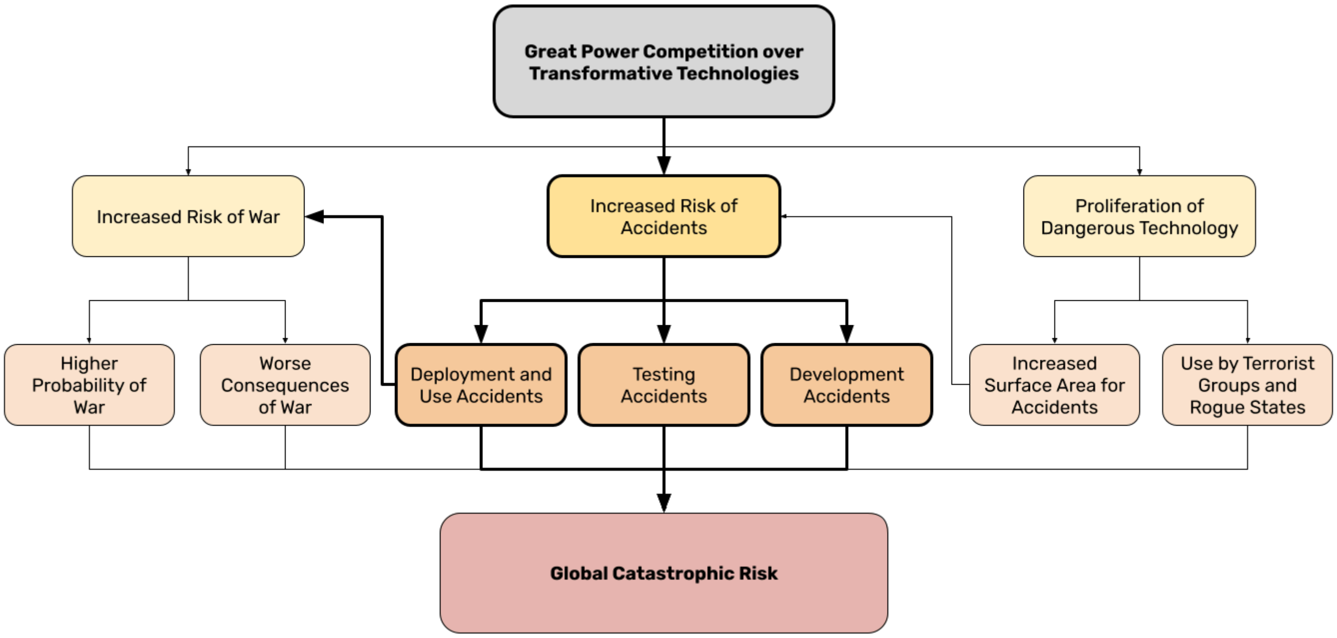
That said, we do venture some tentative conclusions: that more intense technology competition probably makes war somewhat more likely and somewhat more destructive.

It is important to note that we are not just trying to assess how technological development affects the risk of war. We are instead comparing a world with more intense technology competition to a world with less intense technological competition. This is important because the strongest considerations we raise above relate to the intensity of competition. For example, intense competition seems likely to make war somewhat more likely in large part because in a world of intense competition, rival states will probably have taken some hostile actions, like instituting export controls, that have harmed trust and left them with fewer non-violent options in the future. As a result, we've forecasted that a great power war is one to two percentage points more likely in a world with very intense technology competition than it is in a world with mild competition.

More intense technology competition is also likely to accelerate the development of new weapons by driving increased R&D spending and cause states to build larger arsenals. If one thinks future military technologies will emphasize precision over lethality, or defense over offense, this could actually make future wars less destructive — though this does not necessarily follow. However, if one expects the longer-term trend towards increased destructiveness to prevail, or if one expects transformative technologies like cheap biotechnology tools and AI could cause major damage, this acceleration could make wars more destructive. Overall, we believe that competition is likely to exacerbate the long-run trend towards more destructive wars, outweighing the more recent emphasis on precision.

Technology accidents

Unfortunately, a purposeful war is not the only way technology competition could drive existential risk. It could also make catastrophic accidents more likely.



We consider the risk of accidents at several stages of a technology's life cycle:

- **Development accidents** — Accidents and unintended effects may occur while a transformative technologies is being developed (e.g. a laboratory leak of a newly-developed synthetic organism)
- **Testing accidents** — Accidents and unintended effects may occur when such technologies are tested, especially during "field tests" (e.g. unexpected electro-magnetic pulse effects and fallout during nuclear testing)
- **Deployment accidents** — Accidents may occur when such technologies are deployed, i.e. moved into position (e.g. unexpected behavior by an autonomous vehicle)
- **Use accidents** — Accidents may occur when such technologies are used (e.g. mutation of a contagious biological weapon after its use)

There is some evidence that military accidents occur relatively frequently. For example, Danzig notes that during World War II the United States rapidly incorporated new airplane capabilities and sustained large amounts of accidents; on the continental United States alone 15,000 people died in flight accidents.[51] In other words, a full 4% of total US fatalities in WWII arose from accidents unrelated to combat. More recently, during the Cold War, the United States and the Soviet Union also pursued new technologies such as bioweapons despite a high risk of accidents.

This problem of secrecy is not unique to governments. Consider advanced private AI labs, which also operate under high secrecy such that few people understand the actual technologies behind contemporary large language models. Moreover, private citizens can file FOIA requests to learn more about classified government programs, a recourse that is not generally available with private

[51] Danzig, *Technology Roulette,* 8.

companies.[52] Nonetheless, there are several factors that make classified government work more concerning than secretive private R&D:

- The government can monitor and regulate the work of private companies, especially if that work appears to threaten national security, but the reverse is not true
- Public exchange of information can help identify flaws and overconfidence[53]
- Government laboratories may work on systems whose existence is not even known privately
- An apparent tendency towards over-classification can make it difficult for one part of the government technology innovation complex to fully understand and evaluate the work of any other part[54]
- The consequences of leaking classified information are more severe than the consequences of breaching private non-disclosure agreements

Crucially, **development and testing** accidents may occur even when militaries have no immediate plans to deploy or use the technology under development. Soviet research into bioweapons during the Cold War, for example, involved handling and cultivating dangerous pathogens. In 1979 one of these pathogens, anthrax, leaked from a lab and killed dozens of citizens (see Case Study below). Although such research was banned by the Biological Weapons Convention in 1972, Soviet bioweapons research continued until the end of the Cold War. It is difficult to measure compliance in this area and it is possible that bioweapons research continues in some form today. Expert assessments of similar research in other areas, such as the chance of containment failures in high-security BSL-4 labs, suggest that the chance of a containment failure is relatively high (see "Laboratory Accident Rates" on p. 42-43 of Global Catastrophic Biological Risks).[55]

Development and testing accidents are not confined to bioweapons research. Tests of new nuclear weapons, too, have sometimes gone awry. The 1954 Castle Bravo test, for example, contaminated the Marshall Islands with nuclear fallout and afflicted the crew of a Japanese fishing vessel with radiation sickness.[56]

In the future, we might expect tests of new bioweapons, explosives, automated weapons, or other military technologies to pose similar challenges. Militaries will want to test these weapons to learn more about them before they are deployed and used. But even these tests will raise risks of their own. This risk may be exacerbated for artificial intelligence, where increasing the scale of models

---

[52] Thanks to Rosella Capella Zielinski for this point.
[53] Ibid.
[54] Federation of American Scientists, "What is Over-Classification?" https://fas.org/blogs/secrecy/2013/10/overclass/#:~:text=Overclassification%20refers%20to%20the%20classification,precludes%20the%20possibility%20of%20public.
[55] For an attempt to quantify the existential risk posed by these and other biosecurity concerns, see Piers Millett and Andrew Snyder-Beattie, "Existential Risk and Cost-Effective Biosecurity," *Health Security* 15, no. 4 (2017): 373–83, https://doi.org/10.1089/hs.2017.0028.
[56] Toby Ord, Rafaela Hillerbrand, and Anders Sandberg, "Probing the Improbable: Methodological Challenges for Risks with Low Probabilities and High Stakes," Journal of Risk Research 13, no. 2 (March 1, 2010): 191–205, https://doi.org/10.1080/13669870903126267.

may lead to discontinuous advances in capabilities and emergent behavior — including the emergence of dangerous behavior and the risk of loss of control.

The risk of accidents during **deployment and use** is even clearer. The real-world behavior of novel technologies is sometimes difficult to predict. This is especially true for artificial intelligence, where even with rigorous testing and evaluation, performance in a test environment is not necessarily indicative of performance in the world. Some highly advanced AI-enabled systems may even exhibit behavior during testing that looks like "scheming," which further complicates the problem of safe deployment and use.[57]

For military technologies, battlefields are chaotic environments. New weapon technologies often have surprising effects. Machine guns, for example, changed completely the nature of warfare; the resulting trench war of World War I was different, and deadlier, than the 19th-century wars that preceded it. As weapons become more powerful, these unexpected effects could become catastrophic. Future weapons seem to have the potential to cause damage without human operators: consider automated weapons like pilotless drones, bioweapons that can spread uncontrollably. Such technologies are particularly worrying from a global catastrophic risk perspective.

Technology competition seems very likely to *raise* the risks of such accidents. First, more intense competition drives militaries to pursue more powerful technologies that may give them a relative advantage — including technologies that are inherently risky. Second, it pushes militaries to speed up their pursuit of these technologies for fear of their rival getting them first (e.g. rapidly scaling and deploying novel AI systems without rigorous testing and evaluation, for fear of falling behind in a "race"). This will likely come at the expense of some safety checks — a "race to the bottom" on safety. And third, it lowers trust between rival states, making it harder to share information and reach agreements to mitigate these dynamics.

A critical question is how strong these effects are. How much higher is the risk of accidents in a world of intense technology competition than it is in a world of lower tensions? This is an important question for further research, which we do not attempt to answer in this report.

Case Studies: Russian Biological Weapons Development

The history of the Cold War is full of examples of accidents and near-misses involving transformative technologies, driven by military competition and a race to develop new capabilities. The Soviet Union's biological weapons program, despite the evident risks of laboratory leaks, continued, leading to incidents like the 1979 Sverdlovsk Anthrax leak, in which a bioweapons plant

[57] Joe Carlsmith, "Scheming AIs: Will AIs Fake Alignment during Training in Order to Get Power?" (arXiv, November 27, 2023), https://doi.org/10.48550/arXiv.2311.08379.

leaked deadly Anthrax that poisoned the residents of the town of Sverdlovsk.[58] The Soviet Union was also experimenting with deadly pathogens like smallpox and plague in its Biopreparat program; a leak with such weaponized agents could have been catastrophic.

Accidents and war

Another possibility to consider is that an accident may not directly cause a catastrophe, but could lead to inadvertent escalation that ends in a catastrophic war — the arrow pointing to “increased risk of war” in our pathways diagram.

This risk seems likely to grow as the role of technology in threat assessment and military decision-making grows. During the Cold war, for example, technological malfunctions seem to have raised the chance of war breaking out on several occasions. In the 1980s alone, for example:[59]

- A US early warning sensor interpreted a Soviet missile launched during a training exercise as headed for the United States
- A faulty computer chip led to repeated warnings about incoming missiles in US command centers
- A Soviet satellite falsely reported that the US had launched five missiles towards the Soviet Union

In each case, the technical malfunction was discovered before actual retaliation could be launched. However, it seems there was some chance of real escalation. Since the stakes are so high, military decision-making protocols have multiple safeguards, redundant systems, and opportunities for human intervention. This means the chance that any given incident escalates may be quite low. But unless one thinks that chance is zero, then the total risk accumulates over time.

This risk seems highest when new technologies have been recently deployed. For example, the risk of an accidental nuclear war may have gone down over time as early warning technologies like satellites have become more reliable.[60] But the coming decades could see automation continue to spread throughout military systems. Integrating AI could allow for militaries to respond to new information much more quickly, creating a competitive pressure to adopt AI techniques. This could

[58] For an overview of this accident, see David Hoffman, *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy*, 1st edition (New York, NY: Anchor, 2010).
[59] “Accidental Nuclear War: A Timeline of Close Calls,” *Future of Life Institute*, accessed January 12, 2024, https://futureoflife.org/resource/nuclear-close-calls-a-timeline/.
[60] Peter Wildeford, “The Chance of Accidental Nuclear War Has Been Going Down,” accessed January 12, 2024, https://forum.effectivealtruism.org/posts/woBYNgqjgvryF6aav/the-chance-of-accidental-nuclear-war-has-been-going-down.

increase the risk of "artificial escalation" or dangerous dynamics leading to "flash war" in a state of "hyperwar" (the Western term) or "battlefield singularity" (the Chinese term)[61]

On the other hand, by gathering and analyzing more data, AI systems could be safer than conventional early warning systems. For this reason the effect of technology competition, which we expect to accelerate the adoption of such technologies, on the risk of accidental war is somewhat more ambiguous. For a deeper discussion of the effects of autonomy and AI on war, see Founders Pledge's report on Autonomous Weapon Systems and Military AI.

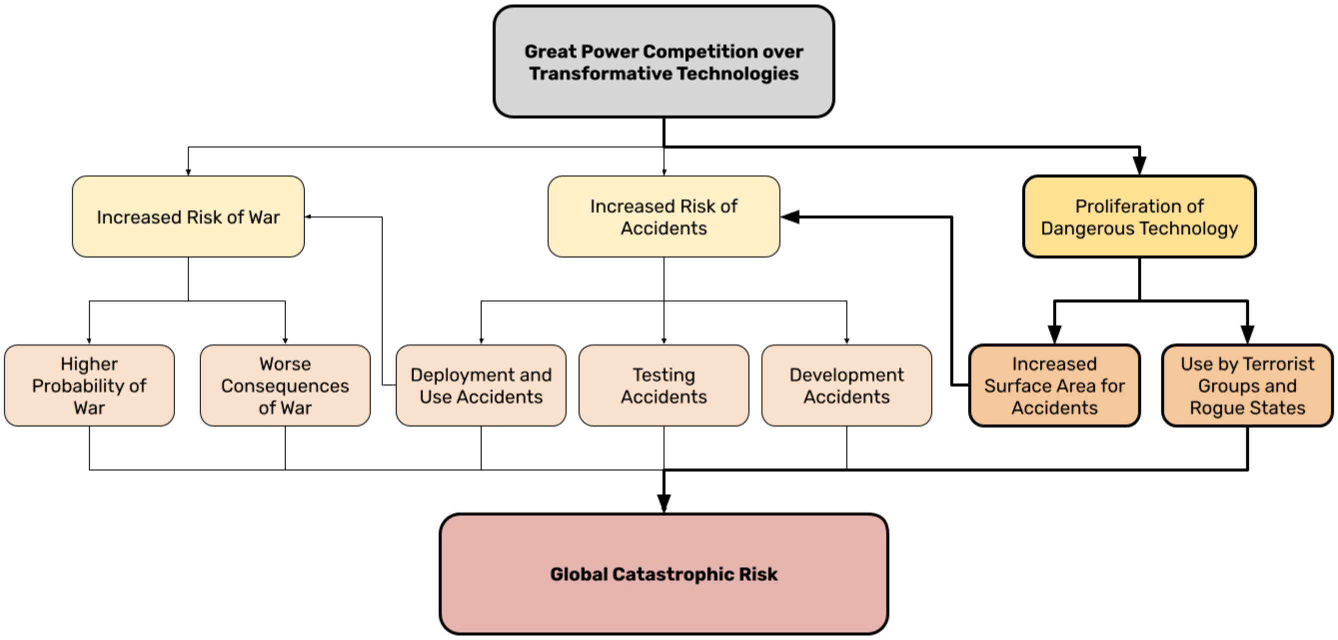
Will technology competition drive the proliferation of transformative technologies?

The last threat pathway is proliferation. Major militaries themselves may have no incentives to conduct mass-casualty attacks and may have stringent safety procedures in place. But the technologies they invent will probably proliferate: over time, more actors will obtain access to them. This can raise risks in three ways:

1. **General Diffusion**: First, the more a technology proliferates, the higher the risk of accidents becomes. For example, if multiple small and medium-sized states open new laboratories to study a synthetic biology technology developed by the United States and China, each laboratory adds to the risk of catastrophic lab leaks.[62] This happens even if all the states in possession of the technology are relatively safety-conscious. As long as there is some level of risk for each actor, then total risk will be higher when more actors are involved.
2. **Diffusion to Less Competent or Cautious Actors**: Second, as powerful technology diffuses, it is more likely to fall into the hands of less competent or less safety-minded actors. Thus, even in situations where technologies *can* be safely managed, the risk of an accident grows because it becomes more likely that one of the actors possessing them will fail to implement adequate safeguards.
3. **Diffusion to Malevolent Actors**: Finally, technology could fall into the hands of groups who want to commit mass casualty attacks. Terrorist groups and doomsday cults, for example, may seek to misuse transformative technologies to create as much damage as possible. And some states may seek to use powerful technologies to pursue geopolitical goals, raising the risk of deployment or use accidents.

[61] Vincent Boulanin *et al.*, "Artificial Intelligence, Strategic Stability and Nuclear Risk" (SIPRI, June 2020), https://www.sipri.org/publications/2020/policy-reports/artificial-intelligence-strategic-stability-and-nuclear-risk.

[62] "Technology proliferation increases technology risks even if others were as invested in safety as we are. These risks multiply as the number of actors and the number of their interactions grow." Danzig, *Technology Roulette,* 9.



Technology competition, again, seems likely to raise proliferation risks. It is likely to accelerate the development of powerful technologies, leading to counterfactually faster proliferation. Faster proliferation may leave less time for states to develop safety technologies and processes or governance frameworks that can mitigate the risks of proliferation. For example, World War II greatly sped up efforts to develop atomic bombs. Their transformative power was obvious, and various schemes for international governance of nuclear weapons and materials were discussed.[63] However, mutual distrust between the USSR and the US undermined these plans, leading to the Cold War arms race, a decades-long threat of nuclear war, and the proliferation of nuclear weapons to seven additional countries so far.

Great power competition over powerful technologies may also raise the salience of these technologies for other actors in the international system, including actors who may be less security-minded or concerned about nonproliferation than the great powers. Depending on the barriers to development, this may lead to the proliferation of dangerous technologies as many well-resourced groups pursue them simultaneously.

The dynamics of great power competition may increase proliferation via theft and spying. For example, recent work in AI has focused on the problem of securing model weights.[64] Well-resourced nation states have cyber tools and intelligence capabilities that are not available to other actors, e.g. corporations. Thus, for example, authors of a recent RAND report on AI model weights emphasize that the most challenging security level for AI labs is “attack vectors available to most high-priority,

---

[63] MacAskill, *What We Owe the Future*, 41-2

[64] See, for example, ongoing work at RAND, “Securing Artificial Intelligence Model Weights.” Sella Nevo et al., “Securing Artificial Intelligence Model Weights: Interim Report” (RAND Corporation, October 31, 2023), https://www.rand.org/pubs/working_papers/WRA2849-1.html.

high-investment attacks by the top cyber-capable state actors."[65] The more that such actors — including the most powerful states — focus these resources on intelligence collection around transformative technologies, the more we would expect such technologies to proliferate.

That said, we are more confident that technology competition speeds the development of new technologies than we are that it causes the invention of technologies that otherwise would have remained undiscovered. Separately, the securitization of technologies may actually lead to improved information security practices, decreasing the risk of proliferation. Again, it is difficult to ascertain how these effects interact, and the magnitude of the effect remains an important question for further investigation.

Case Study: EternalBlue, WannaCry and NotPetya

EternalBlue was a cyber exploit reportedly developed by the U.S. National Security Agency (NSA) to take advantage of a vulnerability in Microsoft's Windows operating system, which was subsequently leaked by a hacker group, and was used to build the technology behind the harmful WannaCry and NotPetya ransomware tools deployed by groups affiliated with Russia and North Korea. This ultimately enabled the 2017 cyber attack on Ukrainian infrastructure. The case illustrates how technology developed by state actors can proliferate and be misused. The consequences of WannaCry and NotPetya were limited and mostly economic, but malware that targets critical infrastructure or AI-enabled cyber tools may have more far-reaching and catastrophic consequences.

Neglectedness

We have argued that competition over transformative technologies is an important problem. While geopolitical trends are hard to predict, we also argued that competition seems more likely to intensify than abate in the coming decades. More intense competition seems likely to accelerate the development of high-risk technologies like bioweapons and advanced artificial intelligence systems. And accelerating such technologies seems likely to make war somewhat riskier, make catastrophic accidents more likely, and lead to wider proliferation.

However, establishing that a problem is important does not necessarily mean that there are good philanthropic opportunities to help solve it. Especially when we lack clear cost-effectiveness data, we also seek to understand how *neglected* the problem is. Neglectedness refers to how much attention, funding, and resources are directed towards a problem relative to its significance. In this

[65] Sella Nevo et al., "Securing Artificial Intelligence Model Weights: Interim Report" (RAND Corporation, October 31, 2023), https://www.rand.org/pubs/working_papers/WRA2849-1.html.

section we attempt to assess the neglectedness of risks from competition over transformative technologies. This is a challenging task, and we intend this section to be a first attempt at defining the problem, rather than a definitive statement of neglectedness.

We first face the challenge of quantifying neglectedness. First, it is difficult to create clear estimates on the relative inattention or lack of funding of a multi-faceted problem like technology competition. Second, there will be a lack of public data on many relevant programs — for example, classified research on major risks from technological investments.

Moreover, there is one major sense in which great power competition is far from neglected; there are large-scale philanthropic efforts designed to **encourage** competition. For example, the Special Competitive Studies Project, a subsidiary of the Eric & Wendy Schmidt Fund for Strategic Innovation, is a major effort to ensure American competitiveness in technology competition. Similarly, many defense-focused think tanks have programs on American competitiveness in key enabling technologies.

In this report, we limit the scope of our analysis of neglectedness to efforts to **understand and mitigate the risks** of great power competition over transformative technologies, such as:

- Research projects to understand risk
- Policy initiatives to translate research into effective policies
- Diplomacy programs focused on reaching international agreements to limit risks

To begin to build an estimate, we searched relevant funding databases — like thePeace and Security Funding Index — to get a sense of how much philanthropic funding goes to initiatives related to technology competition. These contain little evidence of funding for great power competition over high-risk technologies (though the absence of evidence is not evidence of absence). The most relevant grants (included below) appear to come from Open Philanthropy:

Funder	Recipient	Date	Amount	Stated Purpose
Open Philanthropy	CSIS	2020	$118,307	"To explore possible projects related to AI accident risk in the context of technology competition"
Open Philanthropy	CISAC	2020	$67,000	"To explore possible projects related to AI accident risk in the context of technology competition"
Open Philanthropy	Rice, Hadley, Gates and Manuel LLC	2020	$25,000	"Open Philanthropy recommended a contract of $25,000 with Rice, Hadley, Gates and Manuel LLC to explore possible projects related to AI accident risk and technology competition. This falls within our focus area of potential risks from advanced artificial intelligence and was supported through a contractor agreement. While we typically do not publish pages for contractor agreements, we occasionally opt to do so."

Many other think tanks and government research agencies conduct relevant work, some of which is cited in this report, and is not captured in grants databases. Additionally, we are aware of fewer than five dialogues between the great powers focused specifically on strategic competition over key technologies (like AI).[66] If we estimate that we are under-counting such dialogues by 25% (e.g. because we are not privy to all ongoing discussions), and that each of these dialogues costs about $250,000 per year, then a rough point estimate of current spending on diplomatic dialogues is ~$1.9 million. Note that this spending includes both private philanthropy and government support, as agencies like the Defense Threat Reduction Agency (DTRA) are known to support track 1.5 and track 2 dialogues on technologies of concern. This suggests that our lower bound on spending on this issue should be around $2 million a year.

Despite our high uncertainty on this issue, we have decided to provide a rough quantitative estimate of neglectedness. Based on our understanding of the field and searching through relevant funding databases, we are 60% confident that combined government and philanthropic spending targeted on mitigating the risks of great power competition over transformative technologies is between $2 million and $10 million a year (not counting efforts designed to *increase* one side's competitiveness). This estimate is highly sensitive to definitions; we're taking a relatively narrow definition that focuses on major risks exclusively. This is also intended merely as a first step towards quantifying neglectedness. For comparison, $10 million a year is about ⅓ of the annual philanthropic spending on nuclear issues.[67]

Tractability

This section examines the tractability of the problem — the extent to which the problem can be feasibly addressed with the resources available. At first blush, this kind of technology competition seems highly intractable; shaping the international security environment or the trajectory of great power politics is a daunting task for philanthropists. On further investigation, however, there are several reasons to suppose that the issue could be tractable, including that reducing the risks of military-technological innovation is aligned with the interests of major militaries. Overall, tractability remains the biggest open question for this problem — we simply do not know whether philanthropists can affect the diffuse causes or alter the entrenched incentive structures.

Incentives and complexity make this a tough problem

There are two broad categories of reasons to think this is not a particularly tractable problem. First, there are many powerful actors who are strongly incentivized to pursue military superiority and to

[66] These include the Pacific Forum's U.S.-China dialogues on strategic nuclear issues, INHR's AI dialogues, Brookings-Tsinghua AI dialogues, some work by the Carnegie Endowment for International Peace, and others.
[67] It is difficult to compare the number to spending on AI safety, as we believe these numbers have changed dramatically in the last two years, and are not confident in our estimates.

stoke competition over transformative technologies. Second, there are limited opportunities to change, shape, or counterbalance those incentives.

But on specific issues, progress seems possible

We aim to be realistic about making progress on the overall problem of technology competition. Philanthropy-funded diplomacy initiatives or public campaigns may be able to affect overall levels of tension. However, they are likely to be one influence among many. A more promising approach, though, may be to focus on specific issues where those interests are more aligned, but which are still relatively broad and therefore leveraged. These could include:

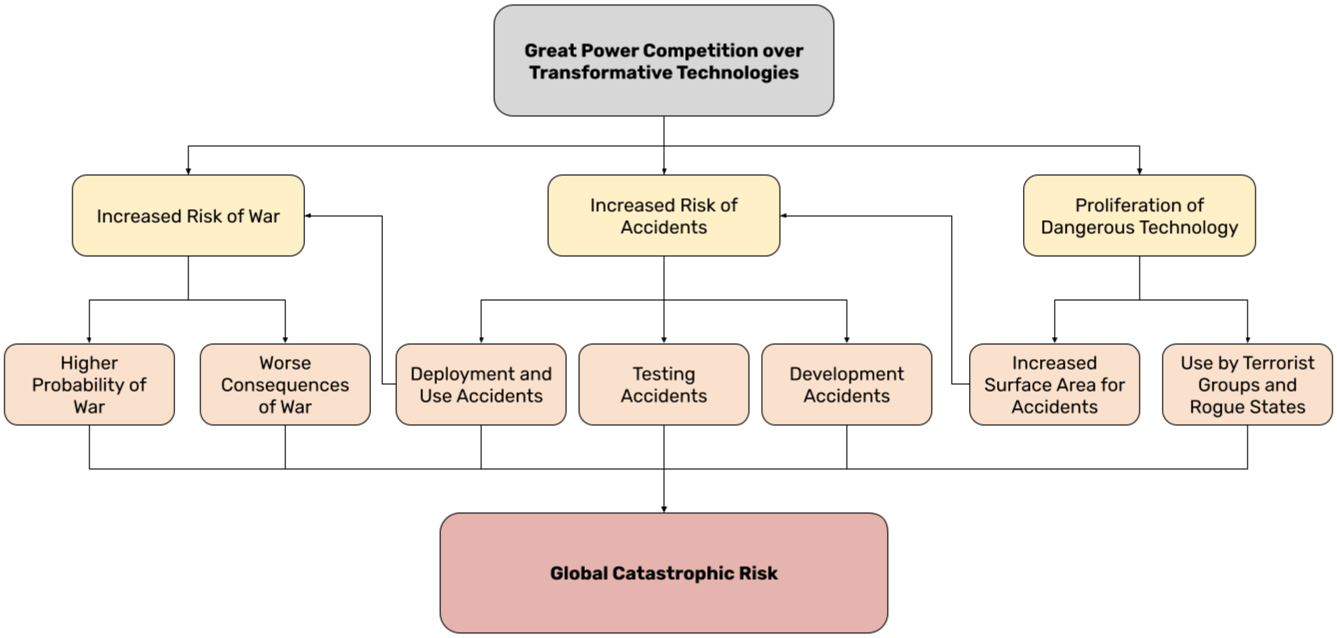
- Improving threat assessment
- Reaching agreements on collective action problems which involve enormous (including existential) risks
- Preventing accidents
    - Militaries want their systems to perform predictably and reliably to avoid accidents and losing control of their weapons.
- Reducing proliferation
    - Militaries don't want weaker states or terrorist groups to get access to transformative technologies
- Avoiding zero- or negative-sum competition, especially when it is exceptionally expensive[68]

The following section dives deeper into the question of how philanthropists could tackle part of this problem via the mechanism of the security dilemma.

Interventions: Causal model

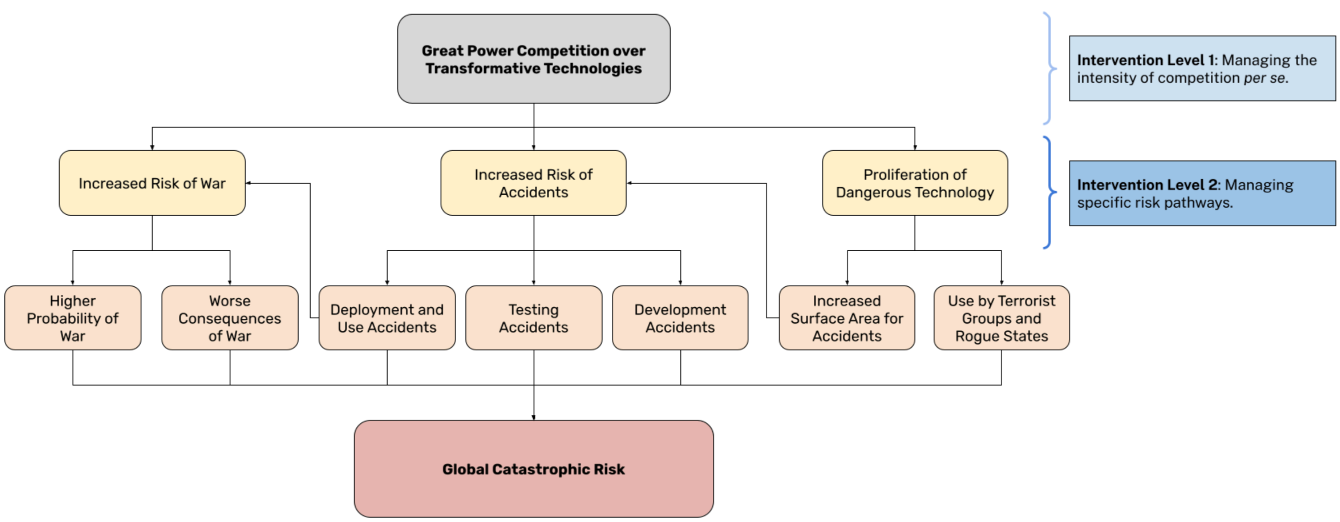
With a better understanding of the importance, neglectedness, and tractability of the problem, we can now return to our causal pathways model to identify the most promising intervention points:

[68] "A suboptimal buildup/race […] might simply result in wasted resources, but not a reduction in security." (Charles Glaser, "When Are Arms Races Dangerous? Rational versus Suboptimal Arming," *International Security* 28, no. 4 (2004): 47, footnote 8.)



Broadly, this suggests two levels of interventions:

1. **Managing the intensity of competition** — Attempting to shape the international security environment, threat perception, threat assessment, and government policies that drive great power competition.
2. **Managing specific risks of competition** —Interventions that seek to e.g. reduce the proliferation of dangerous technologies and minimize the risk of accidents.



Sample Interventions

The following non-exhaustive list is intended to provide a first cut at listing sample interventions to mitigate the risk of great power competition over transformative technologies. Our aim is to spark discussion about potential interventions.

Intervention Level	Sample Interventions
1 – Managing the Intensity of Competition	• **Diplomatic Dialogues** — track 1.5 or 2 dialogues on transformative technologies to build trust and understanding. • **Improved threat assessment** — research and advocacy for improved threat assessment (such as integrating probabilistic forecasting into the intelligence process). • **Policy advocacy for unilateral restraint** — projects to advocate for caution around unconstrained competition over transformative technologies.
2a — Managing Specific Risks of Competition (Risk of War)	• **Crisis communication "hotlines"** — see Call Me, Maybe? Hotlines and Global Catastrophic Risks. • **Confidence-building measures (CBMs)** — promoting rules of the road and shared understanding to reduce the risk of inadvertent escalation. • **Advocacy for formal arms control** — treaty-based arms control.
2b — Managing Specific Risks of Competition (Risk of Accidents)	• **Incident sharing measures** — tools and mechanisms, modeled on e.g. the National Nuclear Risk Reduction Centers, to share incidents, e.g. about AI-related accidents. • **Liability and insurance levers** — using liability and insurance to help price in the risk of accidents with transformative technologies. • **Policy advocacy for "culture of safety"** — using policy advocacy to help push for a culture of safety that includes rigorous assessments of accident risks.
2c — Managing Specific Risks of Competition (Proliferation Risks)	• **Information security** — tools to help prevent the proliferation of technology and information related to transformative technologies, including projects like Securing Artificial Intelligence Model Weights. • **Advocacy around non-proliferation agreements** —Where non-proliferation is in the shared interest of multiple actors, advocacy to pursue formal or informal agreements. • **Counter-terrorism tools, export controls, etc.** — Countering the proliferation of transformative

	technologies using established counter-terrorism tools, export controls, and similar measures.

Preliminary Thoughts on Sample Interventions

Rather than offer a fully-developed framework for evaluating such interventions, we offer some preliminary thoughts for how to think about the impact of interventions designed to reduce the risks stemming from great power competition. Several of these may ultimately lead to "impact multipliers" — general heuristics that can help to guide philanthropists towards high-impact interventions in high uncertainty contexts.

Prioritize All-Hazards Interventions

For the reasons discussed throughout this report, an effective intervention on high-risk technology competition ought to be risk-general or threat agnostic as much as possible. This is intuitive: given the choice between an intervention that is applicable to only one potential risk (e.g. policy advocacy for greater focus on safety on autonomous weapons) and an intervention that is applicable to a wide variety of risks (e.g. policy advocacy for a greater focus on safety for military development and applications in general), philanthropists ought to focus on the latter.

Prioritize Robustness to Technological Change

Additionally, rather than try to predict technological progress, philanthropists ought to invest in funding opportunities that are robust to the changing landscape. The list of "transformative technologies" may seem obvious now, but there are likely lurking unknown unknowns. This principle of robustness for long-term policy planning is explained neatly in a RAND report on Shaping the Next 100 Years: "The goal is to discover near-term policy options that are robust over a wide range of futures when assessed with a wide range of values. Robust strategies will often be adaptive — that is, they will be explicitly designed to evolve over time in response to new information."[69]

Leverage Societal Resources via Policy Advocacy

As discussed in Founders Pledge's reports on climate, nuclear war, and biological risks, philanthropists can generally multiply their impact by focusing on affecting the allocation of much larger societal resources rather than trying to create change directly. This is especially relevant for issues of national security, where private actors simply have less visibility into relevant questions and problems.

[69] Robert J. Lempert, Steven W. Popper, and Steven C. Bankes, Shaping the next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis (Santa Monica, CA: RAND, 2003), 7.

Minimize Downside Risks

Finally, grantmakers in this space ought to attempt as much as possible to minimize potential downside risks. These downside risks include stoking security dilemmas and creating or disseminating information hazards. Defensive interventions that are designed to decrease risk, for example, may be viewed by low-trust parties as offensive. Similarly, interventions may accidentally disclose dangerous information, draw attention to vulnerabilities, or otherwise create information hazards. For more information on how grantmakers can avoid these dilemmas, see Founders Pledge's recent report on Global Catastrophic Biological Risks: A Guide for Philanthropists.

Countering proliferation may be a high-leverage intervention

Our risks framework suggests that countering proliferation of transformative technologies may be a high-leverage node for intervening on the total amount of risk. This is because the proliferation of dangerous technologies increases the total surface area for accidents, which in turn may increase the risk of war. Moreover, even adversary states may have shared interests in countering the proliferation of dangerous technologies, especially to non-state actors like terrorist groups. Others have also recognized the importance of non-proliferation for reducing risks, including:

- 80,000 Hours: "securing the most advanced AI systems may be among the highest-impact work you could do"[70]
- RAND's project on Securing Artificial Intelligence Model Weights[71]

Conclusion and Next Steps

This report has developed a more rigorous framework for analyzing the effects of great power competition over transformative technologies on global catastrophic risks. We identified seven specific effects of this competition in three categories:

1. **Proliferation of transformative technologies** and associated risks
2. **Accidents with transformative technologies** and associated risks
3. **Increased risk of war** including both the probability and consequences of this risk

After examining the effect of competition on each pathway, we have argued that more intense great power competition likely raises total existential risk. The size of this effect is very uncertain; we have argued it is probably significant, but have not attempted to quantify it here.

---

[70] "Information Security in High-Impact Areas: Career Review," 80,000 Hours, accessed January 12, 2024, https://80000hours.org/career-reviews/information-security/.
[71] Sella Nevo et al., "Securing Artificial Intelligence Model Weights: Interim Report" (RAND Corporation, October 31, 2023), https://www.rand.org/pubs/working_papers/WRA2849-1.html.

This framework in turn points to several possible interventions. Despite the large scale of the problem and its seeming intractability, we argued that some targeted interventions could be effective. But further research is needed to develop rigorous frameworks for evaluating and prioritizing these interventions.

Questions for further research include:

- How can analysts, philanthropists, and policymakers think more rigorously about the trade-offs of risk-mitigating interventions that could negatively affect great power competition?
    - For example, AI chip export controls may appear to be beneficial in the near term, but could have diffuse consequences on U.S.-China competition broadly.
- What, if anything, are the effects of competition over transformative technologies on the probability of war?
- How large is the resource difference between great powers and major private organizations?
    - For example, how much more compute could a state-backed “Manhattan Project for AI” provide over what large multinational corporations can provide to frontier AI labs?
- What are promising examples of policy advocacy for cultures of safety and restraint?

We hope that this report can help to provide a framework for thinking about some of these issues, as well as for philanthropists interested in mitigating these risks.