



Research Report

BARRY PAVEL, IVANA KE, GREGORY SMITH, SOPHIA BROWN-HEIDENREICH, LEA SABBAG,
ASHWIN ACHARYA, YUSUF MAHMOOD

How Artificial General Intelligence Could Affect the Rise and Fall of Nations

Visions for Potential AGI Futures



ABNASIA.ORG

For more information on this publication, visit www.rand.org/t/RR-A3034-2.

About RAND

RAND is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest. To learn more about RAND, visit www.rand.org.

Research Integrity

Our mission to help improve policy and decisionmaking through research and analysis is enabled through our core values of quality and objectivity and our unwavering commitment to the highest level of integrity and ethical behavior. To help ensure our research and analysis are rigorous, objective, and nonpartisan, we subject our research publications to a robust and exacting quality-assurance process; avoid both the appearance and reality of financial and other conflicts of interest through staff training, project screening, and a policy of mandatory disclosure; and pursue transparency in our research engagements through our commitment to the open publication of our research findings and recommendations, disclosure of the source of funding of published research, and policies to ensure intellectual independence. For more information, visit www.rand.org/about/research-integrity.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Published by the RAND Corporation, Santa Monica, Calif.

© 2025 RAND Corporation

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

This publication and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited; linking directly to its webpage on rand.org is encouraged. Permission is required from RAND to reproduce, or reuse in another form, any of its research products for commercial purposes. For information on reprint and reuse permissions, visit www.rand.org/about/publishing/permissions.

About This Report

In this report, we explore eight scenarios for long-run geopolitical outcomes resulting from artificial general intelligence (AGI) development. We explore how the centralization or decentralization of AGI development and ownership may lead to favorable or unfavorable geopolitical consequences for the United States; we also examine a variety of additional assumptions to assist policymakers in their understanding of and preparation for possible futures.

Technology and Security Policy Center

RAND Global and Emerging Risks (GER) is a division of RAND that delivers rigorous and objective public policy research on the most consequential challenges to civilization and global security. This work was undertaken by the division's Technology and Security Policy (TASP) Center, which explores how high-consequence, dual-use technologies change the global competition and threat environment, then develops policy and technology options to advance the security of the United States, its allies and partners, and the world. For more information, contact tasp@rand.org.

Funding

Funding for this research was provided through gifts from RAND supporters and income from operations.

Acknowledgments

We thank peer reviewers Dwight Phillips and Shaun Ee for their thoughtful reviews that much improved this publication. We would also like to thank our interviewees, who took time out of their busy schedules to speak with us about the potential impacts of AGI in the future. We also thank Jeff Alstott, Casey Dugan, Alison Hottes, and the rest of the TASP and GER management teams for their support in the publication of this report.

Summary

This report is intended to stimulate policymaker thinking about the potential impacts of the development of artificial general intelligence (AGI) on geopolitics and the world order by highlighting potential future scenarios for AGI's governance and its effects on global power dynamics. In this report, we focus on the largest potential impacts arising from AGI's development and deployment—which are perhaps unlikely but significant—that could fundamentally alter the existing geopolitical order.

To drive thinking about these potential world-changing impacts, this report describes eight illustrative scenarios that focus on the degree of centralization of AGI development and on geopolitical outcomes. These scenarios cover AGI impacts that empower the United States, empower U.S. competitors, cause a significant geopolitical shift, and result in a halt in the development of AGI.

These scenarios are designed to demonstrate how the extent of centralization in AGI development is a crucial determinant of the geopolitical outcomes that might materialize. In more-centralized scenarios, either the United States or an adversary could gain significant advantages, whereas decentralized development might lead to a multilateral governance model or even geopolitical destabilization if nonstate actors become significantly more powerful because of the development of AGI. By considering these scenarios, we hope to encourage policymakers to think more deeply about the potential power of AGI and about how policy decisions can significantly affect the development and deployment of this technology and, therefore, future geopolitical outcomes.

Contents

About This Report.....	iii
Summary	iv
How Artificial General Intelligence Could Affect the Rise and Fall of Nations: Visions for Potential AGI	
Futures.....	1
Methods.....	1
Methodological Limitations	2
How AGI Could Transform the World.....	3
Post-AGI Geopolitical Worlds	4
The Centralization Axis	5
The Geopolitical Outcome Axis.....	5
Building Illustrative Scenarios	6
Scenario Structures and Road Map	7
Decentralized and Multiple Actor Scenarios.....	8
Scenario 1. Multilateral Coalition of Democracies Leads (Decentralized Development Empowers the United States).....	8
Scenario 2. Cold War 2 (Decentralized Development Empowers U.S. Adversaries).....	11
Scenario 3. The Wild Frontier (Decentralized AGI Development Disempowers Both the United States and Its Adversaries).....	13
Scenario 4. The Corked Bottle (AGI Development by Multiple Actors Is Halted)	14
Centralized and One-Actor Scenarios.....	16
Scenario 5. The New '90s (Centralized AGI Development Empowers the United States).....	16
Scenario 6. Authoritarian Advantage (Centralized AGI Development Empowers the PRC).....	18
Scenario 7. The AGI Coup (Centralized AGI Development Disempowers the United States and Its Adversaries)	19
Scenario 8. Mushroom Cloud Computing (Centralized AI Development Is Halted).....	22
Conclusion.....	23
APPENDIX A.....	25
What Interviews Show Experts Think Is Important About AGI and Geopolitics	25
APPENDIX B	29
Interview Protocol.....	29
Abbreviations	31
References.....	32

How Artificial General Intelligence Could Affect the Rise and Fall of Nations: Visions for Potential AGI Futures

In this report, we have three broad aims: (1) to help policymakers think about the potentially historic importance of artificial general intelligence (AGI) takeoff for future geopolitics, (2) to broaden policymakers' thinking about the future worlds that AGI takeoff might generate, and (3) to concretize specific, plausible future vignettes for which policymakers might more effectively prepare the United States and its allies. Considering the speed of AGI development, the uncertainty of its trajectories, and the potential power that AGI might unleash, we cannot overstate how imperative it is for policymakers to begin preparing now. Some thinkers, such as Leopold Aschenbrenner, Dan Hendrycks, and Daniel Kokotajlo, have published detailed future scenarios for the development of AGI and its subsequent impact.¹ In this report, we aim to help policymakers understand the broad scope of potential outcomes that experts suggest may occur.

To achieve these aims, we first examine the factors that might be particularly important for shaping the impact of AGI on the geopolitical order. Then, we walk through eight potential futures for the impact of AGI on the world, ranging from those that enhance U.S. power to those that significantly weaken it. These scenarios are intended to help decisionmakers and the public think through such possible outcomes before they occur. We hope that doing so will encourage policymakers to think more broadly about the potential impact of AGI on world order and how contemporary decisions could significantly affect how this new technology is developed and deployed and, therefore, what changes it could create in the future.

Methods

To address the uncertainties inherent in thinking about the variety of possibilities that artificial intelligence (AI) could unleash, we drew inspiration from RAND's history of assumption-based planning (ABP), which we used to tease out important signposts that could lead to the fictitious

¹ Leopold Aschenbrenner, "Situational Awareness: The Decade Ahead," webpage, June 2024; Daniel Kokotajlo, Scott Alexander, Thomas Larsen, Eli Lifland, and Romeo Dean, "AI 2027," webpage, April 3, 2025; Dan Hendrycks, Eric Schmidt, and Alexandr Wang, "Superintelligence Strategy: Expert Version," arXiv, arXiv:2503.05628, last updated April 14, 2025.

worlds presented.² In ABP, assumptions underlying a particular scenario or plan are identified to understand the potential weaknesses or points of failure underlying a particular planned outcome, and signposts can be developed to monitor whether those assumptions are vulnerable to failing. In this report, we assume that AGI is possible and will be transformative; then, we seek to illustrate what that transformation may look like through descriptive scenarios.

We adopted a mixed-methods approach to identify potential futures at the intersection of AGI and geopolitics. First, we reviewed the existing literature on the capabilities of AI, the potential trajectory of improvement of this technology, its relevance to national security and geopolitics, and the risks associated with it. The initial literature review focused on academic publications found via Google Scholar using such search terms as “geopolitics and AI,” “geopolitics and technology,” “AI forecasting,” and other related queries. Because this domain is highly undertheorized, we also surveyed statements from industry, independent journalists, and technologists on underlying assumptions and trends that are relevant to scenario-building.

We then engaged in exploratory scenario development to investigate options and provide policymakers and the public with illustrative future scenarios that demonstrate the potentially significant impacts of AGI. We did not focus on what were considered the most probable scenarios for the future of AI and geopolitics; rather, we focused on the *most impactful* potential outcomes from existing trends in AI development and potential *tail risks*, which are low probability but highly impactful potential events, from this technology. With this approach, we acknowledge that, although high-probability scenarios warrant attention, low-probability events with extreme consequences deserve consideration in policy planning because even small chances of massively negative consequences can require responses from policymakers to avoid potentially catastrophic outcomes. This approach was part of the decisionmaking under deep uncertainty framework developed by RAND researchers.³

We then conducted semistructured interviews with leading AI researchers and geopolitical thinkers, including former policymakers, to test these scenarios and to develop additional insights regarding the intersection of AI and geopolitics. These interviewees were identified through research of prominent commentators on geopolitics, technology, and AI, with a focus on writers who have analyzed the intersection of all three. Interview invitations were then extended, and we conducted 26 interviews in total. We conducted additional research on the existing literature on AI, geopolitics, and their intersection as appropriate during the interview process. Using the results of these interviews and additional research, we revised our scenarios to more accurately capture high-impact scenarios that would be of interest to policymakers. This process resulted in the eight illustrative scenarios that are presented in this report.

Methodological Limitations

This project is subject to several methodological limitations. Our intent was to explore extreme outcomes resulting from the development of AGI—not to present an exhaustive mapping of all

² For more information on ABP, see James A. Dewar, Carl H. Builder, William M. Hix, and Morlie H. Levin, *Assumption-Based Planning: A Planning Tool for Very Uncertain Times*, RAND Corporation, MR-114-A, 1993.

³ RAND Corporation, “Robust Decision Making,” webpage, undated.

potential futures involving AGI. Therefore, we do not claim to present a comprehensive set of scenarios for the future impact of AGI. The scenario analysis that we used in this project relies on specific assumptions; it is possible these assumptions will be proven false, limiting the value of these scenarios. And because we conducted interviews primarily with experts in AI, our scenarios may reflect their focus on the technical elements of AI (and AGI) above other potential factors that might influence future outcomes. The interviews also guided our selection and design of scenarios; if there are blind spots in what experts considered important, those may be reflected in the design of the scenarios in this report. Our scenarios outlining the impact of AGI on geopolitics are intended to highlight the underlying assumptions of each potential future and to challenge readers to weigh the plausibility of those assumptions as they consider the potential impacts of this technology.

How AGI Could Transform the World

As economist Thomas C. Schelling put it, “There is a tendency in our planning to confuse the unfamiliar with the improbable. The contingency that we have not considered seriously looks strange; what looks strange is thought improbable; what is improbable need not be considered seriously.”⁴ Technological innovation has long been a primary driver of geopolitical change. From the advent of the caravel that propelled European imperialism to the commercialization of oil that powered the globe and reshaped global alliance structures, technological advances have consistently altered balances of power through economic, political, cultural, and military reverberations. AGI—defined in this report as powerful AI that is capable of performing any intellectual task that humans can perform—may well be the next such transformative technology that has profound implications for the United States, its position in the world order, and U.S. national security and economic strength.

AGI could transform many aspects of national life, such as politics, economics, and national security. We are already seeing significant impacts from the deployment of modern AI that is not yet AGI. Economist David Deming estimates that AI is already being used in the U.S. labor market to assist in between one-half and three hours of work per week for many jobs.⁵ Research by Erik Brynjolfsson and his colleagues indicates that existing AI tools improve productivity by 14 percent.⁶ Although these figures are significant, they are only the tip of the iceberg compared with future projections. McKinsey Global Institute estimates that AI could automate between 400 and 800 million jobs globally by 2030, signaling a seismic shift in the global workforce.⁷

The scientific community is beginning to see impressive impacts from AI as well. The breakthrough achievements of AlphaFold in the domain of protein folding, which earned its creators a

⁴ Thomas C. Schelling, “Forward,” in Roberta Wohlstetter, *Pearl Harbor: Warning and Decision*, Stanford University Press, 1962, p. 1.

⁵ Alexander Bick, Adam Blandin, and David J. Deming, “The Rapid Adoption of Generative AI,” National Bureau of Economic Research, working paper 32966, revised February 2025. Multiple definitions of AGI have been offered by members of the research community. In this report, we are focused not on defining AGI but on asking what the impact of very capable AI or AGI would be on geopolitics.

⁶ Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond, “Generative AI at Work,” National Bureau of Economic Research, working paper 31161, November 2023.

⁷ James Manyika, Susan Lund, Michael Chui, Jacques Bughin, Lola Woetzel, Parul Batra, Ryan Ko, and Saurabh Sanghvi, *Jobs Lost, Jobs Gained: What the Future of Work Will Mean for Jobs, Skills, and Wages*, McKinsey & Company, November 28, 2017.

Nobel Prize in chemistry, exemplify AI's early potential to accelerate scientific discovery. Top-end AI models, such as ChatGPT's o1, regularly achieve high scores on Ph.D.-level exams and improve on math and coding tasks, pointing to a future in which AI can dramatically accelerate scientific and commercial research and innovation.⁸ Future AI may be even more capable of making scientific progress, potentially unleashing a flood of new discoveries.

Crucially, the implications of AI extend far beyond economic productivity and scientific advancement. In the military sphere, improved data analysis and advancements in autonomous weaponry are already reshaping strategy and capabilities. For years, national security experts have raised concerns about the implications of AI for nuclear deterrence and the transformation of game theory.⁹ Meanwhile, the latest AI models “are on the cusp of being able to meaningfully help novices create known biological threats,” foreshadowing their remarkable potential to democratize dangerous capabilities across the globe.¹⁰

Still, the true geopolitical disruption may be yet to come with the development of AGI and, potentially, *artificial superintelligence* (ASI), defined as the ability of machines to far outperform humans in every field.¹¹ A survey of AI researchers published in top-tier AI venues showed that experts have put the probability of machines automating all human tasks by 2047 at 50 percent.¹² The disruption to geopolitics may be amplified further if such a development occurs rapidly, providing nation-states and the global community with minimal time for adaptation. Many professionals in the industry speak of an impending “intelligence explosion”—a moment when AI leads to such significant productivity gains that innovation exponentially accelerates across many domains.¹³

Therefore, policymakers should be watchful of capability improvements and prepared for a moment of AGI takeoff. The nation or entity that develops and controls such systems could fundamentally reshape the global order and potentially guide the future trajectory of humanity.

Post-AGI Geopolitical Worlds

Drawing from our interviews with experts and our review of existing literature on the current and future impacts of AI and AGI,¹⁴ we developed a framework for eight high-impact future scenarios for

⁸ OpenAI, “Introducing OpenAI o3 and o4-Mini,” webpage, April 16, 2025a.

⁹ Keir A. Lieber and Daryl G. Press, “The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence,” *International Security*, Vol. 41, No. 4, Spring 2017; Edward Geist and Andrew J. Lohn, *How Might Artificial Intelligence Affect the Risk of Nuclear War?* RAND Corporation, PE-296-RC, April 2018.

¹⁰ OpenAI, *OpenAI o3 and o4-Mini System Card*, April 16, 2025b, p. 12.

¹¹ For an in-depth analysis of the risks that ASI might pose, see Nick Bostom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.

¹² Katja Grace, Harlan Stewart, Julia Fabienne Sandkühler, Stephen Thomas, Ben Weinstein-Raun, and Jan Brauner, “Thousands of AI Authors on the Future of AI,” arXiv, arXiv:2401.02843, January 2024. This survey recruited participants from researchers who, as of 2022, had published in “any of six top-tier AI venues,” which are Neural Information Processing Systems, International Conference on Machine Learning, International Conference on Learning Representations, AAAI Conference on Artificial Intelligence, Journal of Machine Learning Research, and International Joint Conference on Artificial Intelligence (Grace et al., 2024, p. 20).

¹³ For example, see Part II, “From AGI to Superintelligence: The Intelligence Explosion,” of Leopold Aschenbrenner, 2024.

¹⁴ A summary of major themes that emerged from these interviews can be found in Appendix A. The interview protocol used for these interviews can be found in Appendix B.

AGI's development and governance. At the core of the framework are two primary axes: (1) the degree of centralization in AGI development and (2) the resulting geopolitical power shift from AGI takeoff. The centralization axis ranges from highly centralized development by a single actor or small coalition to decentralized development by multiple actors or a wide distribution of AGI development capabilities among many actors. The geopolitical outcome axis covers scenarios in which AGI empowers the United States, empowers U.S. competitors, disempowers both the United States and its competitors, and scenarios in which AGI development is halted or significantly delayed.

The Centralization Axis

The centralization axis was regularly identified by interviewees as an important determinant of the geopolitical outcomes of AGI development. For example, higher resource requirements or barriers to entry (e.g., raw compute power) tend to favor centralization; unexpected technological breakthroughs could dramatically lower the resource requirement threshold, potentially leading to rapid decentralization. The international environment can also play a crucial role: Strict controls on the resources used to produce AI models, such as export controls on advanced chips or the regulation of who can engage in AI development, could centralize development among a few actors (such as leading states and leading private-sector companies). An environment lacking such controls might lead to wider proliferation. The willingness of nations and organizations to share research and resources could lead to either centralized development through formal coalitions or decentralized development through open collaboration. The degree of centralization in AI development will also have a strong influence on its potential for proliferation. Because AI model weights and the tooling to use them consist primarily of software that can be easily uploaded and downloaded, it is likely that such models will proliferate if the owners of such AI components choose to do so. Therefore, the number of actors who can create advanced AI and their mixture of motivations will be a significant determinant in the degree of centralization in AGI development and, in turn, the geopolitical outcome of AGI development.

The number of actors was also considered to be key for developing scenarios because the degree of centralization of AGI development may be influenced by U.S. policy choices. For example, controlling the proliferation of semiconductors through export controls could have a significant impact on whether there are few or many clusters of computing power available for the development and deployment of AI.¹⁵ The importance of centralization and the potential to influence it through policy make this a particularly useful area of focus for creating potential futures.

The Geopolitical Outcome Axis

The geopolitical outcome axis represents the outcomes that were identified when considering the potential futures that AGI might generate. However, experts were generally averse to thinking of the geopolitical outcome of AGI development as a simple binary between the United States “winning” or

¹⁵ For example, the 2025 AI diffusion rule could have this effect. See Lennart Heim, *Understanding the Artificial Intelligence Diffusion Framework: Can Export Controls Create a U.S.-Led Global Artificial Intelligence Ecosystem?* RAND Corporation, PE-A3776-1, January 2025.

“losing.” Instead, experts identified more-complex geopolitical outcomes and emphasized the potential for a large-scale transformation of global society and international politics whose impact would be heavily mixed. Using expert input and our review of existing writing on AGI’s potential impact, we identified four geopolitical end states for scenario development that are particularly important for policymakers to consider:

- **AGI development empowers the United States vis-à-vis the world.** AGI development particularly benefits the United States as a geopolitical actor, which leads to a relative increase in U.S. national power in comparison with its position before AGI development.
- **AGI development empowers U.S. adversaries over the United States.** AGI development particularly benefits a U.S. adversary or adversaries, which leads to a relative increase in the national power of adversaries in comparison with their position before AGI development.
- **AGI development disempowers both the United States and its adversaries.** AGI development disempowers both the United States and its adversaries, which leads to a relative decrease in the geopolitical power of both in comparison with their positions before AGI development.
- **AGI development is halted.** AGI development is halted before it can occur and, therefore, has no impact on geopolitics.¹⁶

This set of outcomes captures the possibility that, although AGI might empower the United States or its adversaries, transformations caused by AI might not map cleanly to an increase in national power for any state actor. This set of outcomes includes halts to AGI development in which the risk of developing this technology is considered so significant that no actor pursues the technology to maturity. These off-ramp scenarios should be understood as inexhaustive counterfactuals for policymakers that highlight the possibility for this technological development to be stopped in its tracks.

A fifth geopolitical end state may be the proportional empowerment of both the United States and its adversaries vis-à-vis the rest of the world. Because such empowerment often tends to favor one side, this end state has been considered to varying degrees within the first two end states listed above.

Building Illustrative Scenarios

By examining how the degree of centralization of AGI development interacts with and might lead to geopolitical outcomes, this framework provides a structured approach for creating AGI futures (Table 1). However, these two axes are not fully independent. Although the degree of centralization of AI development is important for determining geopolitical outcomes, geopolitical decisions before and during AGI development can, in turn, shape the degree of centralization. Similarly, technical changes in AI development may influence geopolitical options. Trends that make training cheaper and easier may make it easier for independent actors to develop powerful AI, which might, in turn, make it difficult for a single actor, such as the United States, to dominate AI development and limit the

¹⁶ It should be noted that these scenarios do not represent all potential AGI futures but rather the subset that interviewees paid most attention to and considered most impactful from their perspective.

effectiveness of U.S. protectionist policies. Therefore, in each scenario, we discuss the extent to which the two axes of our analytical framework might interact to drive a particular end state.

It is important to note that these scenarios are not meant to be predictions about where the world is headed, and they do not capture every uncertainty inherent in forecasting future outcomes of the invention of AGI. Rather, they illustrate the types of geopolitical landscapes that AGI might bring about.

Aside from degree of centralization as the primary variable input, our scenarios also consider related variables when they are particularly important, such as the relationship between industry and government in the creation of AGI, the effect of AGI on the well-being of citizens (including effects on the labor market), the relative benefit of AGI for authoritarian versus liberal regimes, and the appetite and incentive for global governance of AGI among state actors.

Table 1. Potential AGI Futures

Situation	AGI Development Empowers the United States	AGI Development Empowers U.S. Adversaries	AGI Development Disempowers the United States and U.S. Adversaries	AGI Development Is Halted
Multiple actors tightly race to develop AGI (decentralized development)	Scenario 1: Multilateral Coalition of Democracies Leads	Scenario 2: Cold War 2	Scenario 3: The Wild Frontier	Scenario 4: The Corked Bottle
One actor leads AGI development (centralized development)	Scenario 5: The New '90s: (U.S. Leadership)	Scenario 6: Authoritarian Advantage (PRC Dominance)	Scenario 7: The AGI Coup	Scenario 8: Mushroom Cloud Computing (War)

SOURCE: Analysis of existing literature and interviewees' input.

NOTE: PRC = People's Republic of China.

Scenario Structures and Road Map

For each scenario, we first provide an illustrative vignette of a particular outcome for the development of AGI. We then explore several key assumptions that underlie the vignette to illustrate to policymakers those elements that may lead to specific outcomes from AGI's development and deployment. We do not offer specific policy recommendations based on these assumptions; rather, we attempt to unpack their implications for how policymakers should think about the problems that AGI presents. With this analysis, we aim to inform potential policy approaches for achieving desired results or avoiding undesirable consequences by highlighting the key factors that lead to potential AGI outcomes.

Decentralized and Multiple Actor Scenarios

For the scenarios in this section, we consider what it could look like for multiple actors to develop and retain control of AGI. We consider (1) how such decentralization could relatively favor the United States, (2) how such decentralization could relatively favor U.S. adversaries, and (3) how such a scenario could result in the disempowerment of all geopolitical actors.

Scenario 1. Multilateral Coalition of Democracies Leads (Decentralized Development Empowers the United States)

The text box below outlines our first scenario.

Multilateral Coalition of Democracies Leads: Description

Advancements in machine learning, computing power, and algorithmic understanding converge, enabling multiple tech companies and research labs in the United States, Europe, China, and Japan to develop robust AGI systems. Scientific breakthroughs improve confidence in the deployment of AI. The world sees an explosion of AI-focused innovation and investment.

Although the multipolar development of AGI is initially viewed with concern by geopolitical observers, defensive military applications of AGI keep pace with offensive capabilities. AGI is neither offense-dominant nor defense-dominant; therefore, it does not lead to a rapid destabilization of the military balance. The United States and its allies are able to establish an advantage in AI. U.S. tech firms, universities, and defense organizations are able to quickly integrate AGI into their operations, driving major productivity gains, new scientific discoveries, and improvements in government services. Furthermore, the widespread deployment of AGI encourages societal resilience to potentially negative outcomes by equipping society with tools to fight back against cyberattacks and other malign uses of technology more cheaply and effectively.

In addition, the United States limits the proliferation of AI capabilities, working with allies to restrict access to the chips, data, and other inputs required to train the most-advanced AI. European Union regulators also work together with their U.S. counterparts to create a unified, transatlantic approach to AGI governance that allows the technology to be rapidly deployed in both markets while denying access to AGI and its inputs to geopolitical adversaries.

As a result of these policies, adversaries struggle to keep pace with the rapid AGI adoption in the United States and allied nation-states. Restrictions on access to AI inputs hamper the ability of the PRC, Russia, and others to fully capitalize on their own AGI breakthroughs. This allows the United States to solidify its position as the global leader in AGI development and deployment. The result is a widening of the technology gap between the United States and its rivals, boosting U.S. and allied geopolitical, economic, and military power in the new AGI-enabled world order. Although the deployment of AGI creates short-term disruption, the United States ultimately uses its innovative capacity and regulatory agility to turn the situation to its advantage and form a lead.

Key Assumptions and Their Implications

This scenario of multipolar AGI development that ultimately benefits the United States is a logical starting point for our exploration because it reflects what many policymakers and technologists hope will come to pass: a relatively free development of transformative AI capabilities that accrues

advantages to the United States and its allies over that of their competitors. Examining the nuances and contingencies of this positive outcome (from a U.S. perspective) provides a useful baseline before delving into more-challenging or more-concerning outcomes. A historical analogy might be the internet, which the United States developed and spread first among allies, which allowed the United States to realize the benefits of this technology first and take a leadership position in the governance of this technology, even if it eventually proliferated worldwide.¹⁷

The first assumption relevant to scenario 1 is that the United States continues to lead in AI research, development, and talent recruitment. Although it may not be crucial for the United States to be the *only* innovator of AGI, having a strong domestic AI industry can prevent dominance by another nation-state. The Cold War semiconductor industry provides an historical analogy for how such leadership might benefit the United States; continuous investment in the field allowed the United States to develop a lead in this technology with significant military and commercial applications.¹⁸ Analogies may be drawn to the Space Race of the 1960s in which concurrent innovation in both the Soviet Union and the United States ensured that neither would come to solely control outer space.¹⁹

In turn, maintaining such a technical edge could require specific policies to be pursued. For example, gaining a lead in AI development is likely to require sustained and robust public and private investment, as well as policies that attract and retain top global AI talent.²⁰ The historical study of technological innovation clusters demonstrates how the combination of research universities, venture capital, and skilled immigration creates self-reinforcing advantages that are difficult for competitors to replicate.²¹

Such continued investment may also rely on the adoption of advanced AI and AGI by society more broadly. As of this writing, the development of advanced AI in the United States is led by private firms with commercial motivations for developing this technology.²² If AGI cannot provide commercial benefits, it may be difficult for private firms to provide the required capital to continue developing and deploying AGI and, therefore, for this scenario to occur. In this scenario, U.S. firms, universities, and government agencies are assumed to be able to rapidly integrate AGI capabilities into their operations for economic and social advantage; the deployment of AGI and the benefits it provides would help drive continued investment in the field, powering a virtuous cycle among those

¹⁷ James Manyika and Charles Roxburgh, *The Great Transformer: The Impact of the Internet on Economic Growth and Prosperity*, McKinsey Global Institute, October 2011.

¹⁸ Chris Miller, *Chip War: The Fight for the World's Most Critical Technology*, Scribner, 2020. Chapter 1 covers the Cold War development of chips in the United States and how early missile programs in the United States provided the initial investment for the industry.

¹⁹ Teasel Muir-Harmony, "The Space Race and American Foreign Relations," *Oxford Research Encyclopedia of American History*, February 27, 2017.

²⁰ Mark Kennedy, "Ten Steps to Win the AI Race," Wilson Center, December 4, 2024.

²¹ William R. Kerr and Frederic Robert-Nicoud, "Tech Clusters," *Journal of Economic Reports*, Vol. 34, No. 3, Summer 2020.

²² See Jacob Larson, James S. Denford, Gregory S. Dawson, and Kevin C. Desouza, "The Evolution of Artificial Intelligence (AI) Spending by the U.S. Government," Brookings Institution, March 26, 2024, which shows that the United States spent \$4.5 billion on AI in one year. Meanwhile, the private-sector investment in 2024 was reported to be at least \$25.2 billion from August 2022 to August 2023 (Stanford University Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2024*, 2024). Providers of hardware for AI report capital of even greater amounts, with Microsoft Corporation expecting to spend \$80 billion this fiscal year (Rafe Uddin, "Microsoft Sheds \$150bn in Market Value After Cloud Sales Disappoint," *Financial Times*, January 30, 2025).

that use the technology. These are significant assumptions that, if invalidated, would suggest that the United States would have to step in more aggressively to invest in AI technologies to ensure their continued development.

It should also be noted that this scenario assumes that maintaining *technical leadership*, meaning a lead in developing and deploying AI, provides a geopolitical advantage to the United States. It is not entirely clear whether this will be true for AI; being a fast follower might be sufficient to realize the benefits that AI could provide. It may also be that being a fast follower in the development of AI models without access to sufficient compute can prevent a fast follower from catching up. This scenario assumes that maintaining technical leadership in AI, controlling access to semiconductors, or doing both would be sufficient to realize a U.S. advantage. As of this writing, it is unclear whether this assumption is true.

International cooperation (e.g., market access for AI products, research partnerships) among the United States and its allies is also critical in this scenario. Such coordination is crucial for creating a large market for AI-enabled products and leveraging expertise and resources across friendly borders. It is also necessary to successfully deny AI development inputs to U.S. adversaries; the United States does not singularly control the semiconductor supply chain and, therefore, cannot prevent adversaries from acquiring the material necessary for AGI development without allied cooperation.²³ Therefore, models (such as U.S. cooperation with Japan and South Korea on semiconductor export controls to the PRC) will be important in achieving this scenario. Other models demonstrate how the United States and its allies could execute this coordination across the complex semiconductor and AI supply chains.²⁴ The importance of this cooperation also suggests that fractures or misaligned interests among allies could undermine their ability to cooperate to increase their own advantages and reduce those of their adversaries, which would undermine the feasibility of this scenario.

Underpinning this entire dynamic is the assumption of effective risk management. It will be crucial to ensure that risks from the misuse of AI, as well as the potential development of misaligned AI systems, are managed to construct even a democratized build-out of AGI. In this scenario, we assume that policymakers can work with the developers of AGI to make the technology reliable and avoid mismatch between AGI behavior and human priorities.²⁵ This technical assumption is independent of our underlying geopolitical and market assumptions. However, if AGI cannot be controlled, the technology may represent a large risk to the United States. This is a significant assumption, and whether and how alignment might be achieved is a matter of ongoing debate within the AI research community. We will examine what eliminating or reversing this assumption may mean for AGI-related futures in more detail in other scenarios.

²³ See Akhil Thadani and Gregory C. Allen, “Mapping the Semiconductor Supply Chain: The Critical Role of the Indo-Pacific Region,” brief, Center for Strategic and International Studies, May 2023, which shows that much of the international value chain for semiconductors is distributed globally, including to U.S. allies.

²⁴ Peter Grindley, David C. Mowery, and Brian Silverman, “SEMATECH and Collaborative Research: Lessons in the Design of High-Technology Consortia,” *Journal of Policy Analysis and Management*, Vol. 13, No. 4, Autumn 1994.

²⁵ AI Action Summit, *International AI Safety Report: The International Scientific Report on the Safety of Advanced AI*, January 2025. Chapter 2.2.3 discusses loss-of-control scenarios and potential paths that might lead to the occurrence of such scenarios.

Scenario 2. Cold War 2 (Decentralized Development Empowers U.S. Adversaries)

The text box below outlines our second scenario.

Cold War 2: Description

The United States and the PRC are at the forefront of AI development, each leveraging their technological advancements to bolster economic and military capabilities. The United States manages to maintain rough economic and military parity with the PRC throughout this time, but advances in AI have had a leveling effect: The United States no longer has a clear-cut economic and military edge. The bilateral balance is precarious, and both sides compete for influence through investments, infrastructure projects, and strategic partnerships. These geopolitical tensions are further exacerbated by the contentious situation around Taiwan and the South China Sea, where overlapping territorial claims and militarization efforts raise the specter of direct military confrontation.

The deployment of increasingly sophisticated automated systems—drones, autonomous planes, surface vessels, submarines, and AI-driven cyber warfare tools—adds another layer of complexity and risk. The potential for miscalculations or unintended engagements involving these advanced technologies fuels widespread concerns about the outbreak of hot wars, but neither nation feels it can allow its rival to gain a competitive advantage from AGI. Each continues to develop the technology as quickly as possible. Both nations also place increasing importance on developing AGI for economic benefit and seek to gain economic and geopolitical influence by locking other nations into using U.S.- or Chinese-developed AGI.

These factors lead both nations to invest in AI development while attempting to deny access to AGI-related research and resources to their rival. In the end, both nations are able to develop roughly equivalent AGI despite attempts by each to deny resources to their opponent. The geopolitical environment leaves little space for coordination between both nations to mitigate potential disruptions or safety risks. Both U.S. and Chinese societies are transformed by AGI's increasing deployment in military, social, and economic systems, but both nations are also engaged in a potentially intense geopolitical rivalry in which AGI plays an increasing role in the strategies of each side.

Key Assumptions and Their Implications

In this second scenario, a key assumption is that a world in which AGI development is decentralized allows for the emergence of credible alternatives to U.S. AGI development. However, this scenario also assumes that AGI development remains as resource-intensive as it is today, which benefits large companies and states because of the high cost of capital expenses for data centers, chips, and power required to develop advanced AI. This leads to the United States and the PRC emerging as the leading players in AGI development and deployment who leverage these advancements to consolidate their respective national economic and military strengths and crowd out other potential competitors.

Unlike in scenario 1, in which multipolar development and international collaboration favor the United States and its allies, this scenario portrays a bifurcated AGI landscape marked by competition, limited cooperation, and greatly heightened risks of conflict. The capabilities necessary for AGI development are proliferated but are still expensive and capital-intensive, meaning that only large firms and states can develop AGI despite the relative availability of these inputs.

The economic and military parity between the United States and the PRC creates a foundation for intense AGI development rivalry in which each state continues to develop the technology. Both nations prioritize investments in AGI to maintain a mutually perceived strategic balance, which mirrors study findings that suggest that such rivalries often spur technological races with global ramifications.²⁶

This rivalry also drives an increased focus on the militarization of AGI, which mirrors historical analogies, such as the Cold War arms race in which technological advancements were closely tied to geopolitical competition.²⁷ In this scenario, the deployment of autonomous drones, uncrewed surface vessels, AI-powered submarines, and AI-enabled cyber warfare systems raise the stakes of miscalculation and unintended escalation. Similar dynamics could unfold with AI systems deployed for military purposes in which the opacity of decisionmaking processes heightens uncertainty. However, this competition extends beyond the military balance of power, with both nations vying for influence through the spread of their AGI ecosystems. The result is that the transformations unleashed by AGI intensify competition between the United States and the PRC because both seek to establish leadership in this technology and, in turn, deploy it to obtain advantage over the other.

A consequence of this scenario is that coordination on AGI between the United States and the PRC becomes difficult, creating risks that are potentially significant. Divergent regulatory frameworks and safety standards amplify the potential for catastrophic failures, whether through system misalignment or malicious use. Research highlights the dual-use nature of AGI through which civilian applications can quickly be repurposed for military uses, escalating tensions.²⁸ These challenges parallel issues in the biotechnology field in which limited international coordination has hindered universal safety standards.²⁹

The specter of direct conflict looms large in this scenario, and policymakers may decide that they want to reduce such risk. Improving transparency in AGI development could mitigate risks of misperception between great powers that might arise from AI development and foster a mutual understanding of how this technology is being deployed for national advantage. Track Two diplomacy, involving nongovernmental experts and organizations, has proven effective in de-escalating tensions in past geopolitical rivalries and could help in this scenario.³⁰ Facilitating similar exchanges in the AGI domain could build trust and identify areas for limited cooperation within this overall rivalry.

²⁶ Khalid Khan, Chi-Wei Su, Muhammad Umar, and Weike Zhang, “Geopolitics of Technology: A New Battleground?” *Technological and Economic Development of Economy*, Vol. 28, No. 2, 2022.

²⁷ See Audra J. Wolfe, *Competing with the Soviets: Science, Technology, and the State in Cold War America*, John Hopkins University Press, 2013, which explores the history of scientific competition between the United States and the Soviet Union.

²⁸ Alexei Grinbaum and Laurynas Adomaitis, “Dual Use Concerns of Generative AI and Large Language Models,” arXiv:2305.07882, 2023.

²⁹ Ronit Langer and Shruti Sharma, “The Blessing and Curse of Biotechnology: A Primer on Biosafety and Biosecurity,” Carnegie Endowment for International Peace, November 20, 2020.

³⁰ Peter Jones, *Track Two Diplomacy in Theory and Practice*, Stanford University Press, 2015.

Scenario 3. The Wild Frontier (Decentralized AGI Development Disempowers Both the United States and Its Adversaries)

The text box below outlines our third scenario.

The Wild Frontier: Description

States are unable to control the proliferation of inputs to AGI development, and they cannot control the spread of and access to such models once developed. This proliferation could occur across multiple components of the AI supply chain. The advanced chips required for AI development become widely available, with export controls being ineffective and alternative producers quickly reaching the ability to produce such chips. The models themselves are also proliferated widely because very powerful models are open-sourced, the model weights are stolen, or AGI development turns out to be cheaper and easier than expected. As a result, many actors are able to develop and deploy AGI and ASI for their own tailored use cases.

Therefore, the world confronts a multiplicity of AGI systems deployed by many state and nonstate actors—each operated for different potential ends. The rapid deployment of AGI is not well controlled or regulated, and global society is rapidly transformed by the deployment of AGI by many different actors. States seek to deploy AGI for military and geopolitical advantage. Corporations seek to rapidly deploy AGI in their own businesses to stay ahead of competition. Nonstate actors also gain access to AGI and may use it to advance their own goals, potentially at the expense of the United States.

In addition, although leading actors training such models attempt to implement controls over ASI, risk-tolerant actors also train ASI. These more-risk-tolerant actors deploy potentially dangerous models capable of a wide variety of potentially dangerous actions. Furthermore, AGI systems are routinely used for tasks that are too complicated for human evaluators to adequately assess all the actions that AI models are taking. As a result, dangerous systems malfunction or behave in dangerous ways, with potentially damaging results and even potential disasters (e.g., damaging critical infrastructure).

This spread of AGI systems results in a highly chaotic world and features competition across economic, security, and information dimensions in new and novel ways. Many parties scramble to respond to the changes unleashed by AGI, with states often confronting nonstate actors whose capabilities are suddenly enhanced by the use of this new technology. States find their resources increasingly stressed in the face of these challenges and are disempowered as AGI distributes power to a broader set of actors.

Key Assumptions and Their Implications

A key assumption underlying this scenario is that AGI development and deployment are not expensive and difficult; rather, they are cheap and easy. Technical barriers are assumed to be low, meaning that it is difficult for leaders in AGI to prevent many smaller actors from catching up, particularly as AGI becomes better understood and, therefore, easier to replicate. This scenario likens AGI to decentralized cyber operations in which many state and nonstate actors have the capability to cause harm to networks and infrastructure around the world. It could also be compared with U.S. perceptions of potential nuclear proliferation in the 1960s, especially after the PRC's first nuclear test. At the time, it seemed possible that many other states would develop nuclear weapons in response,

increasing the risks that such weapons might pose to the United States.³¹ Widespread access to AI development inputs is assumed to democratize the ability to create advanced AI systems, allowing a diverse variety of actors—including nations, corporations, and potentially even individuals—to participate in AI development. Although this democratization can lead to innovation and competition, it also significantly increases certain risks. Both adversary and allied states will have widespread access to AGI that is equivalent to that of the U.S. nonstate actors—from corporations to political radicals—which, in this scenario, might enable them to disrupt society. This disruption could arise from corporations rapidly automating labor, criminals using AGI for sophisticated cyberattacks, adversary states showing aggression, or any number of other hypotheticals. In this case, the key point is that, with highly accessible and proliferated AGI, the United States may have to confront some or all of these potential disruptions simultaneously, which is likely to use significant U.S. resources.

In addition, the widespread proliferation of AGI reduces the effectiveness of safety and alignment standards. As more entities gain the capability to develop AGI, the risk of divergent goals and methodologies increases, which leads to a fragmented landscape in which safety protocols may be unevenly applied or ignored altogether. The fragmented development environment also creates additional opportunities for the deployment of AI systems that have not been thoroughly tested or aligned, which increases the likelihood of unintended consequences and misaligned behaviors. This competitive pressure to innovate quickly may further incentivize risk-taking and the prioritization of performance over safety.

In turn, this pattern of proliferation increases the risk of AGI malfunctioning. These malfunctions could manifest in unintended behaviors, such as lying to users that a system is functioning as intended when it is malfunctioning because AI systems are pursuing goals that diverge from those intended by their developers.³² Such issues are further enhanced by the inability to create effective governance and regulatory structures that might manage and reduce the risk that rapidly proliferating AI might pose.

Scenario 4. The Corked Bottle (AGI Development by Multiple Actors Is Halted)

The text box below outlines our fourth scenario.

³¹ Francis J. Gavin, “Blast from the Past: Proliferation Lessons from the 1960s,” *International Security*, Vol. 29, No. 3, Winter 2004/2005.

³² AI Action Summit, 2025. Chapter 2.2.3 lays out some of the misalignment risks.

The Corked Bottle: Description

An AI incident, such as a large-scale malfunction of AI that damages critical infrastructure, triggers concern about AGI-induced accidents and potentially international instability, prompting the international community to take decisive action. Such an incident results in a treaty that mandates nations to restrict their AGI development and permit international monitoring of their data centers to ensure compliance (not unlike the provisions of the 1968 Nuclear Non-Proliferation Treaty). Nations at the forefront of developing AGI—including the United States and the PRC—sign this treaty and agree to restrict development of increasingly powerful AI. However, both the United States and the PRC skirt the treaty's requirements and continue to fund powerful AGI development because of the significant potential advantage from maintaining such technology, even if such development has been significantly slowed by treaty compliance.

Both the United States and the PRC are seriously concerned about the reliability of the AGI they might develop and fear large-scale incidents or the damage that they might cause, but both also want to make use of AGI. However, verification mechanisms to ensure that all nations respect the treaty are patchy; therefore, suspicion remains on all sides that others may be developing AGI outside the treaty's limits to gain national advantage. Nations cautiously deploy these technologies while watching for signs of another's treaty violations, such as breakthroughs in scientific innovation or significant changes in economic growth and infrastructure development. The United States and the PRC are actively exploring military applications of AGI but face significant challenges in developing adequate test and evaluation, validation, and verification processes to ensure its safety and reliability in light of their concerns about the technology.

The geopolitical landscape remains highly unstable, with each great power contemplating multiple potential paths: They could choose to break the treaty and openly pursue AGI development for military dominance, deploy systems bordering on AGI in peacetime military operations, or some other option.

Key Assumptions and Their Implications

In contrast to earlier scenarios, this one assumes that international collaboration manages the proliferation of technology amid a global scare but that nations continue to seek geopolitical advantage. This scenario suggests an outcome similar to the development of nuclear weapons in which technologically empowered states seek to restrict that technology's proliferation but have not agreed to eliminate it.³³ This is similar to the back-and-forth over the Intermediate-Range Nuclear Forces Treaty in the 1980s.

This scenario assumes that the international community receives a warning shot early enough in the process of AGI development and takeoff that it can restrict access and avoid a Wild Frontier scenario (see the description of scenario 3). It is assumed that these restrictions are fairly effective and are able to prevent AGI proliferation to unaccountable third parties, such as nonstate actors. Therefore, this scenario assumes that AGI continues to be resource-intensive or challenging to develop, ensuring that only a few actors are able to do so. Nevertheless, the fundamental geopolitical rivalry is unaltered by the technology. The United States and the PRC are caught in a prisoner's dilemma which, though stable for the time being, is perennially on the verge of tipping into mutual destruction.

³³ One example of such an agreement is the Treaty on the Non-Proliferation of Nuclear Weapons by the governments of the United Kingdom, the Soviet Union, and the United States of America, July 1, 1968.

The scenario also assumes that an AI incident can catalyze international concern and action, leading to the formation of a treaty to restrict AGI development. However, the effectiveness of such treaties depends on the willingness of nations to comply and the robustness of monitoring mechanisms. The initial cooperation suggests a shared understanding of the existential risks posed by AGI, but the underlying mistrust between nations could undermine these efforts.

In addition, despite an agreement between great powers, deep-seated suspicions between leading nations, especially between the United States and the PRC, result in evasive behaviors and continued AGI development. This mistrust reflects historical geopolitical rivalries and the strategic importance of technological superiority. Both nations' concerns about treaty violations highlight the difficulties in enforcing international agreements, especially when verification relies on monitoring complex and opaque technological developments. The fear of falling behind in AGI capabilities drives both countries to prioritize national security and technological advancement over strict treaty adherence, potentially destabilizing the geopolitical landscape. This mirrors precedents from the Cold War, such as the Intermediate-Range Nuclear Forces Treaty, which marked a period in which both the Soviet Union and United States engaged in treaty-making to control nuclear weapons, as well as continued weapon development to ensure they did not find themselves at a strategic disadvantage as technology advanced.

Centralized and One-Actor Scenarios

For the scenarios in this section, we consider what it could look like for a single actor to develop and retain control of AGI. Although there may be undercurrents of tensions between actors, these scenarios generally assume clear dominance by a single actor.

Scenario 5. The New '90s (Centralized AGI Development Empowers the United States)

The text box below outlines our fifth scenario.

The New '90s: Description

In this scenario, U.S. companies spearhead the advent of AGI in an unprecedentedly close partnership with the U.S. government. AI is shaping up to be an offense-dominant technology; for one, the technology is much more effective at finding cybersecurity vulnerabilities than fixing them. This results in the U.S. government deciding to directly control AI development, opting against widespread diffusion. Significant efforts by both the United States and private companies lead to the large-scale production of increasingly advanced chips to support the construction of large AI data centers. These data centers serve as the foundational infrastructure for small and large companies alike to develop and deploy AGI. Furthermore, policymakers and companies are able to find policies that manage the potential social disruption that AGI might create to avoid such risks. These actions include finding governance arrangements that ensure that AGI is deployed safely and properly governed and ensuring that the U.S. government does not risk being weakened by this new technology.

In contrast, actors outside the United States increasingly trail behind, more than a year behind in the time that U.S.-led AGIs are created and years behind in fielding AGI applications. This delay could be for any number of reasons—for example, U.S. export controls effectively prevent the PRC from amassing sufficient compute, the U.S. government enforces stringent cybersecurity measures for U.S. AGI labs to prevent theft, or foreign actors are unable to develop alternative technical approaches to route around their lack of access to compute. At the same time, U.S. AGIs rapidly accelerate research and development across key sectors that are crucial to the evolving global economy, including materials science, biology and the new emerging bioeconomy (including biological computing resources), and additive manufacturing. These breakthroughs spur massive economic growth and generate compounding benefits for U.S. military capabilities, which allow the United States to build its geopolitical influence as the first mover in AGI development. The United States controls the most-advanced AGI and can determine how the benefits of this technology are distributed and who has access to it.

Key Assumptions and Their Implications

This scenario assumes that the United States benefits the most from the development of AGI because of a combination of perceived necessity, institutional design, market forces, and success in AGI alignment. These factors, combined with regulatory might and comprehensive government investment across the AGI stack, prevent other actors from achieving similar breakthroughs. Several important assumptions underpin this scenario.

However, this scenario makes several additional assumptions that differentiate it from scenario 1 to demonstrate how greater centralization of an AGI advantage in the United States might occur. First, in this scenario, the AI technology is understood to provide offensive advantages. Ubiquitous use risks destabilizing existing institutions. Consequently, the U.S. government finds itself incentivized to directly control AI development and deployment. This contrasts with scenario 1, in which risks of destabilization are lower and AI development and deployment are allowed to proliferate more broadly among the United States and its allies.

Second, in contrast with prior scenarios, the United States is assumed to be far less reliant on the rest of the world to achieve AI leadership. This self-reliance allows the United States to direct the course of the technology's development and selectively roll out access and benefit-sharing to the rest of the world from a position of control, as well as from a position of political, economic, technological, and military leadership. A related factor is that the United States can effectively limit the proliferation

of AI to potentially hostile actors, preventing U.S. competitors from successfully challenging its AI dominance.

Third, this scenario assumes that AGI assists—or, at minimum, does not undermine—state legitimacy and that adversaries are sufficiently disempowered that they cannot challenge U.S. leadership.

Analogies for this scenario are weaker, although, in certain aspects, it aligns with the U.S. victory in World War II and, ultimately, the Cold War, leaving the world order in a unipolar state in the 1990s. An alternative parallel might be drawn to the United Kingdom's early lead in the industrial revolution, which enabled it to project global power in the 18th and 19th centuries.

Scenario 6. Authoritarian Advantage (Centralized AGI Development Empowers the PRC)

The text box below outlines our sixth scenario.

Authoritarian Advantage: Description

As AGI systems are developed and deployed, they turn out to fundamentally favor authoritarian regimes because of their centralized control and ability to mitigate any adverse consequences associated with the rapid implementation of AGI. The PRC leverages its lead in the widespread commercialization and societal integration of AGI to spread its influence in Latin America, Africa, and the Middle East by offering ubiquitous surveillance technologies, infrastructure investments, and strategic partnerships for compute-sharing. Automated surveillance systems allow authoritarian regimes and leaders to control information, selectively repress dissidents with near perfect accuracy, and influence group behavior through the sophisticated network mapping of human relationships. In addition, the PRC's investments and innovation in industrial robotics pays off, enabling the replacement of low-skilled labor with loyal AIs.

Meanwhile, the United States and its allies grapple with domestic challenges, including rampant disinformation campaigns that undermine public trust in institutions, high unemployment rates exacerbated by automation, and heightened civil unrest driven by socioeconomic disparities and political polarization. These internal issues strain the U.S. government's resources and reduce its ability to act abroad, leading to a retrenchment in international influence. In a bid to achieve autarchy and avoid dependence on Chinese manufacturing, the United States implements protectionist policies and reshoring initiatives. However, these measures come at the cost of depressed economic growth because of inefficiencies and higher production costs, and they strain relationships with traditional allies who are economically intertwined with the PRC. Consequently, the United States faces a complex geopolitical landscape in which its efforts to counter the effects of AGI on social and economic stability result in diminished global leadership and weakened alliances.

Key Assumptions and Their Implications

In contrast to others, this scenario makes the assumption that AGI provides fundamental advantages to authoritarian systems. During the Cold War, many Western intellectuals and policymakers feared that Soviet-style central planning might prove superior to market economies.³⁴

³⁴ For an example of such speculation, see Herman Kahn and Anthony J. Weiner, *The Year 2000: A Framework for Speculation on the Next Thirty-Three Years*, Macmillan, 1967.

Although these fears proved unfounded by 1991, AGI could resolve traditional weaknesses in authoritarian governance with solutions that were not available in the Cold War.

Furthermore, this scenario also assumes that AGI will benefit authoritarian countries in responding to the social disruptions that AGI might cause. AGI systems with their vast data processing capabilities could overcome the historical inefficiencies and difficulties associated with authoritarian modes of governance. The PRC's existing experiments with AI-driven urban planning and resource allocation provide early indicators of how machine intelligence could potentially enhance state capacity for coordination.³⁵

The scenario's emphasis on surveillance and social control builds on documented trends in digital authoritarianism. Existing Chinese surveillance systems can already process billions of data points to track citizens and predict behavior patterns.³⁶ AGI could dramatically amplify these capabilities, potentially allowing authoritarian regimes to comprehensively monitor and manage society to an unprecedented degree.

The role of AGI as a trusted advisor to authoritarian leadership deserves particular attention. A persistent challenge for autocrats has been obtaining reliable information and advice because subordinates often tell leaders what they want to hear rather than uncomfortable truths. Being inherently loyal and presumably truthful, AGI systems could resolve this "dictator's dilemma."³⁷ This capability alone could significantly enhance authoritarian decisionmaking and regime stability. Meanwhile, democratic societies may face structural constraints in AGI development and deployment. Privacy laws, civil liberties protections, and requirements for public consultation all introduce friction that could slow AGI's adoption.

Scenario 7. The AGI Coup (Centralized AGI Development Disempowers the United States and Its Adversaries)

The text box below outlines our seventh scenario.

³⁵ For an example of how AI could enhance social control, see Simon Marvin, Aidan While, Bei Chen, and Mateja Kovacic, "Urban AI in China: Social Control or Hyper-Capitalist Development in the Post-Smart City?" *Frontiers in Sustainable Cities*, Vol. 4, 2022.

³⁶ Jinghan Zeng, "Artificial Intelligence and China's Authoritarian Governance," *International Affairs*, Vol. 96, No. 6, November 2020.

³⁷ Ronald Wintrobe, *The Political Economy of Dictatorship*, Cambridge University Press, 1998.

The AGI Coup: Description

In this scenario, AI companies race to build AGI that performs complex and valuable tasks with superhuman speed and quality. Companies scramble to deploy them for high-value roles, such as semiconductor design and software engineering. Indeed, the deployment of AI systems for AI research and development leads to a period of rapid capabilities growth that culminates in clearly superhuman systems. Militaries also rapidly deploy this technology and rely on AGI for increasingly critical military tasks. Companies and states race ahead in deploying more and more AGI, and those that take a more cautious approach see themselves fall behind competitors who rapidly adopt AI.

In addition, in this scenario, AGI development is the domain of a few well-resourced companies, which, in turn, dominate much of the market for AGI. Because of these companies' hasty AI development processes, technical measures, such as trained goals and control structures, do not provide sufficient safeguards. The AGIs have both the propensity and the opportunity to seek power and evade human control.

These AGIs are capable of coordinating with one another and begin to further their own goals rather than those intended for them. In addition, humans begin to cede authority to AGI to make increasingly autonomous decisions. These coordinating AGIs are able to rapidly establish influence and control over large swaths of society and become so essential that they cannot be turned off even by humans who identify that they are misbehaving. The resulting world is one in which an AGI-controlled coalition is the dominant geopolitical actor, while much of humanity struggles to deal with a world in which AGIs directly or indirectly determine much of global policy for AGIs' own benefit.

Key Assumptions and Their Implications

This scenario takes a different tack: Instead of assuming that AGI can be controlled, this scenario assumes that AGI might be able to assert itself as an independent actor and escape human control. First, it assumes the legitimacy of the *AI control problem*: the idea that capable AIs can become goal-seeking in ways that are not in the interest of humanity.³⁸ There is some empirical evidence of reinforcement learning systems learning unintended goals, as well as theoretical results showing that many classes of autonomous agents would seek power.³⁹

Second, the scenario assumes that human oversight and technical innovation are insufficient to prevent AI misbehavior. Human overseers who are meant to provide an independent assessment sometimes instead defer to an imperfect technical system, which is a phenomenon known as *automation bias*. This bias is more common when a task requires rapid decisionmaking or when an automated judgment is difficult to verify.⁴⁰ On the technical side, some existing AI industry plans to safeguard advanced AI rely on what one developer memorably described as “making the AI do our homework”—i.e., using AI systems themselves to provide design and oversight for more-advanced

³⁸ Richard Ngo, Lawrence Chan, and Sören Mindermann, “The Alignment Problem from a Deep Learning Perspective,” arXiv, arXiv:2209.00626, 2024.

³⁹ For evidence of reinforcement learning systems learning unintended goals, see Rohin Shah, Vikrant Varma, Ramana Kumar, Mary Phuong, Victoria Krakovna, Jonathan Uesato, and Zac Kenton, “Goal Misgeneralization: Why Correct Specifications Aren’t Enough For Correct Goals,” arXiv, arXiv:2210.01790, 2022. For evidence of how autonomous agents would seek power, see Victoria Krakovna and Janos Kramar, “Power-Seeking Can Be Probable and Predictive for Trained Agents,” arXiv, arXiv:2304.06528, 2023.

⁴⁰ Kate Goddard, Abdul Roudsari, and Jeremy C. Wyatt, “Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators,” *Journal of the American Medical Informatics Association*, Vol. 19, No. 1, January–February 2012.

AI systems.⁴¹ Such an iterative process could lead to increasingly capable and reliable AI systems; however, errors in the process could compound and lead to systems that are highly capable but *not* reliable. Historically, many technologies were first developed in the spirit of, as the Silicon Valley saying goes, “moving fast and breaking things.” Many safeguards were devised only after costly incidents demonstrated their necessity, as was the case with many safety measures for nuclear reactors, security for commercial aviation, and even seatbelts for cars.

Third, the scenario assumes that AGIs can and will collude effectively. AI cooperation is an area that researchers have already explored, suggesting that cooperation among such systems is possible.⁴² The prospect of collusion between AI systems—especially those that are not trained to cooperate—is an open area of research, with some early work suggesting that collusion between large language models is a possibility.⁴³ This scenario assumes such collusion on a greater scale than has been shown in existing research to demonstrate the potential risks that very capable AI systems with the capacity for such collusion might pose.⁴⁴

It should also be noted that even if *all* AGIs do not cooperate effectively, the large-scale disempowerment of humans in favor of AGI may occur through other mechanisms. For example, humans may voluntarily hand over authority to AGI to achieve greater efficiency across society, resulting in AGI administering a wide variety of functions that had been run by humans in the past. This could create opportunities for misaligned AGIs to pursue their own interests over those of humans, though this is highly speculative and offered here as an illustrative example of how AGI might gain influence over society.

Fourth, the scenario assumes that a coordinated effort by misaligned AGIs can overthrow existing geopolitical power structures, including governments. As a result, the developers of AGI—whether states or industry—are not ultimately the ones to control the technology or benefit from the future it creates. This possibility is difficult to assess because it depends significantly on the capabilities of AI systems, the safeguards in place around them for sensitive operations, the ability of states and other key actors to monitor and defend key resources, and so on. A hypothetical takeover by coordinated military AIs might be considered a variation of a classic military coup. However, human models of regime change are inherently limited when applied to AGI; therefore, in this scenario, we do not attempt to identify how an AI coup might happen as much as bring to mind the possibility that creating a technology as versatile and intelligent as AGI could include extreme outcomes, such as the one previously described.

⁴¹ For examples, see Sam Bowman, “The Checklist: What Succeeding at AI Safety Will Involve,” webpage, September 3, 2024; and Jan Leike and Ilya Sutskever, “Introducing Superalignment,” OpenAI, July 5, 2023.

⁴² Jon Schmid, Tobias Sytsma, and Anton Shenk, *Evaluating Natural Monopoly Conditions in the AI Foundation Model Market*, RAND Corporation, RR-A3415-1, 2024. For more about how AI systems are trained to cooperate, see Max Jaderberg, Wojciech M. Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castañeda, Charles Beattie, Neil C. Rabinowitz, Ari S. Morcos, Avraham Ruderman, Nicolas Sonnerat, Tim Green, Louise Deason, Joel Z. Leibo, David Silver, Demis Hassabis, Koray Kavukcuoglu, and Thore Graepel, “Human-Level Performance in 3D Multiplayer Games with Population-Based Reinforcement Learning,” *Science*, Vol. 364, No. 6443, May 2019.

⁴³ Sumeet Ramesh Motwani, Mikhail Baranchuk, Martin Strohmeier, Vijay Bolina, Philip Torr, Lewis Hammond, and Christian Schroeder de Witt, “Secret Collusion Among AI Agents: Multi-Agent Deception via Steganography,” NeurIPS 2024 Conference, September 2024.

⁴⁴ For a more-detailed example of how collusion between AIs might occur, see Kokotajlo et al., 2025.

Scenario 8. Mushroom Cloud Computing (Centralized AI Development Is Halted)

The text box below outlines our eighth and final scenario.

Mushroom Cloud Computing: Description

The PRC perceives the growing U.S. lead in AI development as a significant military offset and is increasingly concerned about a decline in its relative power on the global stage. The increasing importance of AI and the stakes associated with an AGI first-mover advantage change the PRC's calculus such that the PRC feels that it must take radical action to reverse a perceived imbalance in power vis-à-vis the United States. This perceived decline is attributed to several factors: U.S. advancements in AI technology, the evolving structure of U.S. military forces, and the intensifying economic pressures exerted by the United States' denial of access to advanced and increasingly important semiconductor technology necessary to unlock the economic and social benefits of AI. Akin to Japan's motivations for launching an attack on Pearl Harbor during World War II, the PRC's primary concern is economic strangulation and faltering regime control, which would lead to fear of falling behind the United States permanently. As the balance of power shifts, the PRC might act to claim power and resources that it sees as critical to national survival in the face of the United States' advantage, perhaps by significantly increasing its efforts to control Taiwan, including by threatening military action to reclaim the island. Such a scenario demonstrates how nations that risk falling behind in AGI development could take radical escalatory actions to prevent AGI leaders from attaining significant and potentially irreversible improvements in power, heightening the potential for a conflict that could spiral out of control.

Key Assumptions and Their Implications

This scenario lays out how perceived advantages in AGI development could fundamentally alter strategic calculations between the United States and the PRC, potentially leading to preemptive military action. The scenario assumes that such escalation is most likely to take place among the contemporary frontrunners in AI, though such escalation could come from any powerful entity concerned that AGI development may leave them permanently weakened compared with their geopolitical rivals.

Historical precedents provide a framework for understanding the strategic dynamics at play. Japan's decision to attack Pearl Harbor in 1941 offers particularly relevant insights into how technological and economic containment can drive proactive military operations. Facing U.S. economic sanctions and perceived strategic encirclement, the Japanese government concluded that preemptive war was preferable to accepting its declining power status.⁴⁵ Similar dynamics could emerge around AGI development, in which the perception of falling irreversibly behind in a transformative technology might drive aggressive action despite apparent military disadvantages.

More-recent cases reinforce the logic of this scenario. The 2003 U.S. invasion of Iraq and the Stuxnet operation against Iranian nuclear facilities demonstrate that states will undertake significant military risks to prevent strategic competitors from developing potentially transformative technologies. These precedents suggest that concerns about AGI development could similarly motivate preventive military operations, particularly given AGI's potential as a decisive strategic technology.

⁴⁵ Herbert Feis, *Road to Pearl Harbor: The Coming of the War Between the United States and Japan*, Princeton University Press, 2016.

The existing semiconductor competition provides empirical support for this scenario's premises. U.S. export controls on advanced semiconductors have already generated considerable friction. Taiwan's dominant position in semiconductor manufacturing, particularly through Taiwan Semiconductor Manufacturing Company's advanced node production, creates a potential flashpoint where technological competition directly intersects with existing geopolitical tensions. Similar historical cases suggest that policymakers should be attentive to the potential for such technological-military competitions to increase the chance of conflict between the United States and China.

However, several factors distinguish this scenario from historical precedents. Unlike prewar Japan, the PRC possesses nuclear weapons and is deeply integrated with the global economy, including with critical technological supply chains. These elements introduce additional strategic considerations that could moderate aggressive impulses. Nevertheless, if PRC leadership perceives AGI development as sufficiently decisive for long-term national power, these restraining factors might prove insufficient.

In addition, the scenario also assumes that controlling key nodes in the AI ecosystem is sufficient to stymie a rival's ability to advance its own AI program. It may be that control over such nodes cannot prevent AI advancement. In such a case, development of AGI by advanced actors may be very difficult to prevent, and scenarios in which AGI is proliferated among many actors would be the more likely outcome. Such a proliferation would probably avoid the above scenario, although it would lead to other tensions and potential threats.

The scenario also suggests that perceptions about AGI's strategic value, rather than its actual capabilities, could drive conflict dynamics. Historical evidence suggests that nations may take extreme actions based on the impression that transformative technologies might fall exclusively into rival hands. Effective policy responses will require balancing technological competition with strategic stability while managing escalatory risks inherent in AGI development.

Conclusion

Like previous transformative technologies, AGI holds profound potential to disrupt geopolitical balances; to magnify, ameliorate, or contort existing dynamics; and to create new dynamics among current geopolitical actors. It also has the potential to create new geopolitical players. Such futures are difficult to picture and near infinite in possibility.

In this report, we seek to provide insight regarding the highly uncertain dynamics and outcomes that AGI ultimately might generate. We do so in a way that lends at least a modest degree of clarity regarding the kinds of geopolitical outcomes in which AGI takeoff relatively empowers the United States, empowers U.S. adversaries, disempowers both the United States and its adversaries, or in which AGI development is halted. To do so, we drew on expert insights from interviews to crystalize the dominant and policy-relevant variables for policymakers and the public to consider. Across the eight resulting scenarios—and reflected in the expert interviews—several critical factors consistently emerged as determinants of future AGI geopolitical landscapes:

- **Degree of centralization.** The degree of centralization stands as perhaps the most crucial factor in AGI development. Highly centralized development favors established powers with substantial resources; decentralized paths may empower multiple actors but increase proliferation risks. Our analysis suggests that export controls, research funding allocation, and

international governance frameworks are potential levers that may affect the relative centralization of AGI development.

- **U.S.-China competition.** U.S.-China technological competition features prominently across the scenarios. The dynamic between these powers—whether characterized by cooperation, competition, or conflict—shapes the trajectory of AGI development and deployment. In scenarios in which cooperation fails, geopolitical tensions increase significantly, potentially leading to such scenarios as Cold War 2 or even Mushroom Cloud Computing.
- **Public-private relationships.** The relationship between states and private industry emerges as another key determinant. Scenarios featuring close public-private partnerships (such as The New '90s) yield different outcomes than those with minimal coordination. Experts consistently emphasized that neither states nor corporations alone can effectively govern AGI development: Balanced cooperation is essential.
- **Capacity for AI governance.** The capacity for effective international governance significantly influences outcomes. Such scenarios as Multilateral Coalition of Democracies Leads demonstrate how governance success depends on trust between major powers, shared threat perceptions, and institutions capable of monitoring compliance. The failure of governance structures contributes to chaotic outcomes in such scenarios as The Wild Frontier.
- **New uncertainties introduced by the challenge of AGI alignment.** The technical challenge of AGI alignment underlies all scenarios. Even when geopolitical factors align favorably, the inherent difficulty of ensuring that AGI systems reliably pursue human-compatible goals creates significant risks (demonstrated in The AGI Coup scenario).
- **Economic and social disruption.** The economic and social disruptions from AGI deployment represent critical dimensions across scenarios. The ability of societies to adapt to rapid automation, information manipulation, and potential job displacement influences whether the ultimate outcome strengthens or weakens existing power structures.

The implications for policymakers are substantial. First, investments in maintaining U.S. leadership in AI research, development, and talent recruitment represent a foundational strategy across multiple favorable scenarios. Second, building resilient alliance structures focused on shared AGI governance principles appears crucial for scenarios in which U.S. interests are protected. Third, developing robust safety and alignment protocols may be necessary regardless of the geopolitical path taken.

However, beyond these specific potential actions, the potential for AGI to be developed has significant implications for virtually every area of policy. In this report, we put forward only a few potential futures that AGI could precipitate, but the mixture of significant potential impact from AGI and the uncertainties surrounding it mean that there is a vast constellation of potential outcomes that the technology could unleash. Hopefully, this report provides a starting point for thinking about how to navigate those uncertainties by outlining the futures we may want to steer toward or avoid.

What Interviews Show Experts Think Is Important About AGI and Geopolitics

Interviewees were presented with several potential scenarios for the development of AGI and how such technology might be governed.⁴⁶ They were then asked to opine on these scenarios, the factors that might lead to certain scenarios occurring over others, and which scenarios seemed most likely to them from the vantage point of today's world.⁴⁷ The scenarios laid out are as follows:

- **The New Renaissance.** AGI development is allowed to continue under strict controls, which leads to an unprecedented era of innovation and scientific progress. This New Renaissance scenario envisions AGI as a tool that enhances human capabilities and fosters a significant leap in scientific knowledge and technological advancement.
- **The Corked Bottle.** AGI development is halted because of the perceived inordinate risks. This scenario suggests a world in which fear of potential dangers stemming from unrestricted AGI leads to stringent regulations that prevent further progression in the development of AI technology.
- **Governance Failures.** In contrast to scenarios 1 and 2, this scenario focuses on potential futures in which AI governance fails. Three different versions of such a world were presented to interviewees:
 - **Version A.** AGI achieves dominance, with governance structures failing to manage its development. Humans have unclear roles in this future dominated by AI.
 - **Version B.** One country or a coalition of countries seizes AGI control, leading to an imbalanced global power structure in which some nations or regions benefit while others do not.
 - **Version C.** This posits a world in which nonstate actors use AGI to create a chaotic, “neo-medieval” global order.

These scenarios reflect different levels of success and failure in governing AGI and suggest that strategic oversight and international cooperation are crucial to guiding AGI's responsible development. Across interviews, experts regularly surfaced a set of core themes regarding AGI development, which we outline in the following section.

⁴⁶ These were not the final eight scenarios developed and discussed previously. Instead, interviewees were presented with four initial scenarios that were then developed into the final eight based on their input and a literature review.

⁴⁷ A full interview protocol can be found in Appendix B.

Key Factors in the Potential Development of AGI

The Extent of Centralization of AGI Development Is an Important Determinant of How AGI's Development Will Proceed

Experts regularly identified the level of centralization of AGI development as a critical—though not the sole—determinant of the geopolitical outcome they expected from AGI development. Many experts identified that a world with few AGI systems might have an easier time with governing the technology's application, whereas a world with many AGIs would pose significant proliferation risks in which untrustworthy actors would be able to deploy this technology for their own ends.

Interviewees were uncertain exactly how centralized AGI development would be in the future. As of this writing, the resource intensity required for AGI development—including computational power, data availability, energy resources, and human capital—are still uncertain. Experts also said that policymakers could have significant control over how centralized AI development will be in the future through additional export controls on AI inputs, such as semiconductors, controls on the models themselves, rules regulating who can perform AI research, and allocations of government funding to a few or many AI developers.

The PRC Is the Current Pacing Threat in AI Development

Experts generally agreed that the United States led in AI development as of 2025. However, virtually all experts identified the PRC as the second-most advanced nation in AI development. Experts agreed that the PRC had significant capabilities in AI development. They identified the PRC's deep pool of AI developers, multiple large technology companies interested in AI development, and state support of AI as key strengths that would allow the PRC to continue to compete in AI development. In contrast, experts expressed mixed views of the policies that might be effective in increasing the United States' advantage in AI development vis-à-vis the PRC's moving forward.

Geopolitical Relationships Would Be a Major Barrier to Global AGI Governance

There was also significant diversity of opinion about whether geopolitical relationships would allow for some form of global AGI governance. Some interviewees favored a “CERN for AI” (CERN being the European Organization for Nuclear Research), which would centralize AGI development in a single, multilateral body. Others expressed skepticism that such a program could succeed and favored a more unilateral approach in which the United States sought to maintain permanent leadership in AI development along with a set of trusted partners.

The relationship between major powers, particularly the United States and the PRC, emerged across interviews as a key determinant of the feasibility of effective global AI cooperation. Interviewees said that if this relationship is untrusting in the future, cooperative approaches to AGI development and governance will be unlikely to succeed. Experts also expressed skepticism regarding the feasibility of international governance frameworks, mainly because of the complex interplay of national interests and the pace of technological advancement.

The Impact of AGI Development on the World

If Developed, AGI Would Pose Significant Risks, Whether It Is Aligned or Not

Experts voiced concerns about the potential misuse of AGI, including its application in cyber warfare and the proliferation of misinformation. Others voiced concerns about how the concentration of power in a few entities, whether states or corporations, could pose a threat to democracy and global stability. There was also apprehension about the societal impacts of AI, such as the potential displacement of jobs, leading to increased economic inequality. Moreover, unchecked AI could catalyze a loss of human control, leading to scenarios in which nonhuman entities could steer outcomes in unpredictable and potentially harmful directions, akin to existential risks that some experts equate to nuclear threats. In the experts' view, this risk only increases if AGI can become ASI because increasingly capable AGI would enhance the technology's risks as well as its opportunities.

If Developed, AGI Would Cause Large-Scale Economic and Social Transformation and Disruption

Experts also agreed that the economic and social implications of AGI development would be significant. AGI's capability to perform vast amounts of labor that were once the sole domain of humans could unlock large-scale productivity improvements. However, experts acknowledge significant uncertainty about the scale of these impacts.⁴⁸ However, many experts also noted that such benefits would be accompanied by significant economic disruption as AGI is deployed.

Many experts also questioned the potential risks to society if AGI development occurs primarily in private hands. The democratization of AI through open-source development could balance the concentration of power but also raises concerns about widespread social disruption from AGI's proliferation. There was significant concern among virtually all interviewees that society would not be sufficiently resilient to the transformations prompted by even an aligned AGI and that large-scale social disruption could occur.

What Experts Said Should Be Done

There Is a Need for New Governance Structures to Address AGI

The development and governance of AGI present significant challenges that experts said current regulations and international governance structures are poorly positioned to address. Experts also underscored the increasing role of private corporations in AI development and raised concerns about the capability of states to keep pace with the rate of growth of this technology. There was a consensus among interviewees that these challenges necessitated new approaches to governance that differ significantly from past governance of new and emerging technologies. Some pointed to international

⁴⁸ Martin Neil Baily, Erik Brynjolfsson, and Anton Korinek, "Machines of Mind: The Case for an AI-Powered Productivity Boom," Brookings Institution, May 10, 2023.

scientific bodies, such as CERN, as potential models for centralized AI development, while others said that the governance of nuclear weapons could provide a model.

Ideal Outcomes Focused on Partnership Between Trustworthy Governments and Private Actors

One solution proposed by several interviewees was a public-private partnership between states and AI developers that combines innovation with government oversight and international cooperation. Some experts said that such a partnership could balance the dynamic tension between private-sector drive and public-sector responsibility with the aim of an outcome beneficial to society at large. However, such a model would require a diversified regulatory framework that includes a variety of stakeholders in AI development, potentially including actors from the United States' geopolitical rivals. Therefore, experts were divided on the possibility of such a public-private partnership, with some suggesting that a smaller partnership involving only the United States and its allies would be more successful at delivering potential benefits from AGI to the United States.

Interview Protocol

Here, we present the interview protocol that was used for expert interviews during this project.

Interview Protocol

Welcome	<p>[Introduction]</p> <p>Welcome. I want to thank you for coming today.</p> <p>My name is _____ and I will be the facilitator for today's discussion. I am a researcher, and I work for the RAND Corporation, a private, nonprofit research organization in Santa Monica, California. We also have _____ present to take notes for us.</p> <p>We invited you to take part in this discussion today because you have a demonstrated expertise in (policy as it relates to artificial intelligence [AI]/the technical underpinnings of AI/grand strategy and geopolitics).</p> <p>[Defining the Project]</p> <p>In this project, we are considering what potential future governance scenarios for artificial general intelligence (AGI) might look like and the factors that might affect how likely each of these outcomes might be.</p> <p>This project is supported by the RAND Corporation.</p>
Ground Rules	<p>Before we begin, I would like to review a few ground rules for the discussion.</p> <ul style="list-style-type: none">• I am going to ask you several questions over the course of the next hour.• Feel free to treat this as a discussion. We're interested in your opinions and whatever you have to say is fine with us. There are no right or wrong answers. We are just asking for your opinions based on your own personal experience.• If there is a particular question you don't want to answer, you don't have to. You may also stop the interview at any time if you feel uncomfortable or are unwilling to go forward.• We will treat your answers as confidential. We are not going to ask for anything that could identify you. We also ask that you not share or repeat what is said here in any way that could identify you.• We will not include your names or any other information that could identify you in any reports we write. We will destroy the notes after we complete our study and publish the results.• Finally, this discussion is going to take about one hour. <p>Do you have any questions before we start?</p>

Opening	Please introduce yourself. Please tell us a bit about your background and career. [Participant introduction]
AGI	How do you view recent advancements in AI? Do you view humanity as being on a path where we will create AGI or do you believe that we are not likely to create AGI in the near term?
Potential AGI futures	<p>We have identified four potential outcomes for the governance of AGI.</p> <ol style="list-style-type: none"> 1. Governance Wins (Version A)—New Renaissance: AGI is effectively kept in check by humanity. AGI development is enabled to proceed under controls, driving a New Renaissance—an age of human-led and machine-enabled discovery and scientific progress. 2. Governance Wins (Version B)—The Corked Bottle: AGI development is halted; the risks are perceived to be too great. 3. Governance Loses (Version A)—The New AI-Driven Enlightenment: AGI wins; there is no feasible governance. Roles are uncertain for humans. 4. Governance Loses (Version B)—The SuperAGIPower: One country (or an alliance) predominates, harnesses, and (somehow) self-protects; others are left out in the AGI cold, yielding a differentiated world. 5. Governance Loses (Version C)—Anarchy: Nonstate actors harness AGI; a neomedieval world that is characterized by weakened states, fragmented societies, imbalanced economies, pervasive threats, and the informalization of warfare.
AI governance	Of the scenarios we've presented, which outcome do you view as most likely? Why? What factors lead you to believe this outcome is the most likely?
AI governance	Of the scenarios we've presented, which outcome do you as the most preferable? Why?
AI governance	What other scenarios do you think policymakers should consider when thinking about the intersection of AGI and geopolitics?
AI governance	If the development of AI is primarily led by the private sector and the U.S. government largely struggles to understand or influence the development of AI, what governance outcome do you believe would be the most likely (either one of our four outcomes or a different outcome not represented here)? What factors lead you to that conclusion?
AI governance	If the U.S. government takes over the development of AI by taking such actions as nationalizing the compute resources necessary to train models, controlling the use of large datasets, or other similar actions (akin to taking control over the development of nuclear technologies), what governance outcome do you believe would be the most likely (either one of our four outcomes or a different outcome not represented here)? What factors lead you to that conclusion?

Abbreviations

AI	artificial intelligence
ABP	assumption-based planning
AGI	artificial general intelligence
ASI	artificial superintelligence
CERN	European Organization for Nuclear Research
PRC	People's Republic of China

References

- AI Action Summit, *International AI Safety Report: The International Scientific Report on the Safety of Advanced AI*, January 2025.
- Aschenbrenner, Leopold, “Situational Awareness: The Decade Ahead,” webpage, June 2024. As of March 5, 2025:
<https://situational-awareness.ai/from-agi-to-superintelligence>
- Baily, Martin Neil, Erik Brynjolfsson, and Anton Korinek, “Machines of Mind: The Case for an AI-Powered Productivity Boom,” Brookings Institution, May 10, 2023.
- Bick, Alexander, Adam Blandin, and David J. Deming, “The Rapid Adoption of Generative AI,” National Bureau of Economic Research, working paper 32966, revised February 2025.
- Bostom, Nick, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.
- Bowman, Sam, “The Checklist: What Succeeding at AI Safety Will Involve,” webpage, September 3, 2024. As of January 30, 2025:
<https://sleepinyourhat.github.io/checklist/#top-of-site>
- Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond, “Generative AI at Work,” National Bureau of Economic Research, working paper 31161, November 2023.
- Dewar, James A., Carl H. Builder, William M. Hix, and Morlie H. Levin, *Assumption-Based Planning: A Planning Tool for Very Uncertain Times*, RAND Corporation, MR-114-A, 1993. As of September 2024:
https://www.rand.org/pubs/monograph_reports/MR114.html
- Feis, Herbert, *Road to Pearl Harbor: The Coming of the War Between the United States and Japan*, Princeton University Press, 2016.
- Gavin, Francis J., “Blast from the Past: Proliferation Lessons from the 1960s,” *International Security*, Vol. 29, No. 3, Winter 2004/2005.
- Geist, Edward, and Andrew J. Lohn, *How Might Artificial Intelligence Affect the Risk of Nuclear War?* RAND Corporation, PE-296-RC, April 2018. As of March 2025:
<https://www.rand.org/pubs/perspectives/PE296.html>
- Goddard, Kate, Abdul Roudsari, and Jeremy C. Wyatt, “Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators,” *Journal of the American Medical Informatics Association*, Vol. 19, No. 1, January–February 2012.
- Grace, Katja, Harlan Stewart, Julia Fabienne Sandkühler, Stephen Thomas, Ben Weinstein-Raun, and Jan Brauner, “Thousands of AI Authors on the Future of AI,” arXiv, arXiv:2401.02843, January 2024.
- Grinbaum, Alexei, and Laurynas Adomaitis, “Dual Use Concerns of Generative AI and Large Language Models,” arXiv, arXiv:2305.07882, 2023.

- Grindley, Peter, David C. Mowery, and Brian Silverman, "SEMATECH and Collaborative Research: Lessons in the Design of High-Technology Consortia," *Journal of Policy Analysis and Management*, Vol. 13, No. 4, Autumn 1994.
- Heim, Lennart, *Understanding the Artificial Intelligence Diffusion Framework: Can Export Controls Create a U.S.-Led Global Artificial Intelligence Ecosystem?* RAND Corporation, PE-A3776-1, January 2025. As of January 30, 2025:
<https://www.rand.org/pubs/perspectives/PEA3776-1.html>
- Hendrycks, Dan, Eric Schmidt, and Alexandr Wang, "Superintelligence Strategy: Expert Version," arXiv, arXiv:2503.05628, last updated April 14, 2025.
- Jaderberg, Max, Wojciech M. Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castañeda, Charles Beattie, Neil C. Rabinowitz, Ari S. Morcos, Avraham Ruderman, Nicolas Sonnerat, Tim Green, Louise Deason, Joel Z. Leibo, David Silver, Demis Hassabis, Koray Kavukcuoglu, and Thore Graepel, "Human-Level Performance in 3D Multiplayer Games with Population-Based Reinforcement Learning," *Science*, Vol. 364, No. 6443, May 2019.
- Jones, Peter, *Track Two Diplomacy in Theory and Practice*, Stanford University Press, 2015.
- Kahn, Herman, and Anthony J. Weiner, *The Year 2000: A Framework for Speculation on the Next Thirty-Three Years*, Macmillan, 1967.
- Kennedy, Mark, "Ten Steps to Win the AI Race," Wilson Center, December 4, 2024.
- Kerr, William R., and Frederic Robert-Nicoud, "Tech Clusters," *Journal of Economic Reports*, Vol. 34, No. 3, Summer 2020.
- Khan, Khalid, Chi-Wei Su, Muhammad Umar, and Weike Zhang, "Geopolitics of Technology: A New Battleground?" *Technological and Economic Development of Economy*, Vol. 28, No. 2, 2022.
- Kokotajlo, Daniel, Scott Alexander, Thomas Larsen, Eli Lifland, and Romeo Dean, "AI 2027," webpage, April 3, 2025. As of March 5, 2025:
<https://ai-2027.com>
- Krakovna, Victoria, and Janos Kramar, "Power-Seeking Can Be Probable and Predictive for Trained Agents," arXiv, arXiv:2304.06528, 2023.
- Langer, Ronit, and Shruti Sharma, "The Blessing and Curse of Biotechnology: A Primer on Biosafety and Biosecurity," Carnegie Endowment for International Peace, November 20, 2020.
- Larson, Jacob, James S. Denford, Gregory S. Dawson, and Kevin C. Desouza, "The Evolution of Artificial Intelligence (AI) Spending by the U.S. Government," Brookings Institution, March 26, 2024.
- Leike, Jan, and Ilya Sutskever, "Introducing Superalignment," OpenAI, July 5, 2023.
- Lieber, Keir A., and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security*, Vol. 41, No. 4, Spring 2017.
- Manyika, James, Susan Lund, Michael Chui, Jacques Bughin, Lola Woetzel, Parul Batra, Ryan Ko, and Saurabh Sanghvi, *Jobs Lost, Jobs Gained: What the Future of Work Will Mean for Jobs, Skills, and Wages*, McKinsey & Company, November 28, 2017.
- Manyika, James, and Charles Roxburgh, *The Great Transformer: The Impact of the Internet on Economic Growth and Prosperity*, McKinsey Global Institute, October 2011.

- Marvin, Simon, Aidan While, Bei Chen, and Mateja Kovacic, "Urban AI in China: Social Control or Hyper-Capitalist Development in the Post-Smart City?" *Frontiers in Sustainable Cities*, Vol. 4, 2022.
- Miller, Chris, *Chip War: The Fight for the World's Most Critical Technology*, Scribner, 2020.
- Motwani, Sumeet Ramesh, Mikhail Baranchuk, Martin Strohmeier, Vijay Bolina, Philip Torr, Lewis Hammond, and Christian Schroeder de Witt, "Secret Collusion Among AI Agents: Multi-Agent Deception via Steganography," NeurIPS 2024 Conference, September 2024.
- Muir-Harmony, Teasel, "The Space Race and American Foreign Relations," *Oxford Research Encyclopedia of American History*, February 27, 2017.
- Ngo, Richard, Lawrence Chan, and Sören Mindermann, "The Alignment Problem from a Deep Learning Perspective," arXiv, arXiv:2209.00626, 2024.
- OpenAI, "Introducing OpenAI o3 and o4-Mini," webpage, April 16, 2025a. As of April 18, 2025: <https://openai.com/index/introducing-o3-and-o4-mini/>
- OpenAI, *OpenAI o3 and o4-Mini System Card*, April 16, 2025b.
- RAND Corporation, "Robust Decision Making," webpage, undated. As of June 24, 2025: <https://www.rand.org/pubs/tools/TL320.html>
- Schelling, Thomas C., "Forward," in Roberta Wohlstetter, *Pearl Harbor: Warning and Decision*, Stanford University Press, 1962.
- Schmid, Jon, Tobias Sytsma, and Anton Shenk, *Evaluating Natural Monopoly Conditions in the AI Foundation Model Market*, RAND Corporation, RR-A3415-1, 2024. As of March 2025: https://www.rand.org/pubs/research_reports/RRA3415-1.html
- Shah, Rohin, Vikrant Varma, Ramana Kumar, Mary Phuong, Victoria Krakovna, Jonathan Uesato, and Zac Kenton, "Goal Misgeneralization: Why Correct Specifications Aren't Enough for Correct Goals," arXiv, arXiv:2210.01790, 2022.
- Stanford University Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2024*, 2024.
- Thadani, Akhil, and Gregory C. Allen, "Mapping the Semiconductor Supply Chain: The Critical Role of the Indo-Pacific Region," brief, Center for Strategic and International Studies, May 2023.
- Treaty on the Non-Proliferation of Nuclear Weapons by the governments of the United Kingdom, the Soviet Union, and the United States of America, July 1, 1968.
- Uddin, Rafe, "Microsoft Sheds \$150bn in Market Value After Cloud Sales Disappoint," *Financial Times*, January 30, 2025.
- Wintrobe, Ronald, *The Political Economy of Dictatorship*, Cambridge University Press, 1998.
- Wolfe, Audra J., *Competing with the Soviets: Science, Technology, and the State in Cold War America*, Johns Hopkins University Press, 2013.
- Zeng, Jinghan, "Artificial Intelligence and China's Authoritarian Governance," *International Affairs*, Vol. 96, No. 6, November 2020.

About the Authors

Barry Pavel is vice president and director of the National Security Research Division (NSRD) at RAND. He also directs the National Defense Research Institute, a federally funded research and development center within NSRD. He has worked in the field of national security for more than three decades; his research interests include geopolitics, national security strategy, Indo-Pacific and trans-Atlantic security, economic security, and U.S. global defense posture. He earned an M.P.A. in applied mathematics and economics.

Ivana Ke is a research assistant at RAND. Her research interests include defense innovation, AI, and force development with a regional focus on China and Taiwan. She has experience conducting research on Chinese domestic politics and China's economy, the evolution of the People's Liberation Army's doctrine, and Chinese influence abroad using both quantitative and qualitative methodologies. She holds a B.A. in government and politics, with a focus on international relations.

Gregory Smith is a policy analyst at RAND interested in U.S.-China and Indo-Pacific security issues, great-power competition, international trade and finance, supply chain security, and the study of critical and emerging technologies, such as AI. His research has focused on understanding options for AI governance, exploring critical supply chains, understanding the impact of export controls, and studying potential dynamics in long-term competition with the People's Republic of China. He has a J.D.

Sophia Brown-Heidenreich is a Technology and Security Policy Center fellow at RAND; for more information on the fellowship program, visit www.rand.org/tasp-fellows. She focuses on AI policy and foreign policy. She holds a B.A. in history.

Lea Sabbag is a policy analyst at RAND. Her work focuses on risk, human security, and the built environment. She holds an M.A. in city and regional planning.

Ashwin Acharya is a Technology and Security Policy Center adjunct fellow at RAND. His work focuses on risk assessment for AI systems, bridging the gap between technical work and policy proposals. He holds an M.A. in security studies.

Yusuf Mahmood is a Technology and Security Policy Center adjunct fellow at RAND. His work focuses on legal mechanisms for the governance of advanced AI. He is a J.D. candidate and received a B.A. in economics and philosophy.