

Working Paper

JOEL PREDD

Riding the Coming Wave

The Urgent Need for Contingency Plans for AGI Futures

RAND Global and Emerging Risks

WR-A3691-1
June 2025

RAND working papers are intended to share early insights and solicit informal peer review. This working paper has been approved for circulation by RAND but has not been peer reviewed or professionally edited or proofread. Review comments are welcomed at the SocArXiv version of this paper. This working paper can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. This working paper does not necessarily reflect the opinions of RAND's research clients and sponsors. RAND® is a registered trademark. Learn more at www.rand.org.

For more information on this publication, visit **www.rand.org/t/WRA3691-1**.

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.

WR-A3691-1

About This Working Paper

This working paper was originally drafted for discussion at the Geopolitics of Artificial General Intelligence (AGI) Council, a monthly RAND-hosted convening of researchers, private sector representatives, government officials, and others to discuss technological and policy developments pertaining to artificial intelligence (AI). In its current form, this paper is intended for those who participate in the Council and a broader audience of policymakers and national security and technology experts working outside of government.

Based on conversations with the Council and with RAND colleagues over the past year and a half, during which time I have led the Geopolitics of AGI Initiative in RAND, I have arrived at a set of assumptions regarding the U.S. government's ability to produce a strategy for adopting and adapting to transformative artificial intelligence (AI). These assumptions underpin this paper and are as follows: the accelerating pace of technological progress in frontier AI, coupled with inherently limited governmental capacity for anticipatory policymaking, means that a national strategy to adapt to a future with AGI plausibly will emerge from government's responses to events rather than from a top-down analytic or policymaking process. The choices that government officials and business leaders make in response to such events could influence the future just as much, if not more, than institutions or policies intended to govern the development and diffusion of AGI.

Based on that diagnosis, this working paper suggests that analytic and governmental resources should be allocated immediately to

- identifying near-term events that pose acute threats to the status quo *and* represent signposts that a systemic shift in national security, economics, or geopolitics may be imminent;
- developing "break glass" plans for government to respond to those contingency-signposts; and
- developing and exercising capabilities and competencies to execute and adapt those plans as needed.

Embedded within those plans may be courses of action that are infeasible or unwise to pursue amid today's uncertainties and debates, or given the opportunity costs. However, it is likely that such plans would prove useful to the U.S. government if and when transformative AI is reached by enabling policymakers to seize opportunities presented by unfolding events, and to thereby shape futures to ends desirable for the United States and for humanity.

Technology and Security Policy Center

RAND Global and Emerging Risks 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 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

The Geopolitics of AGI Initiative is independently initiated and conducted within the Technology and Security Policy Center using income from operations and gifts from philanthropic supporters, which have been made or recommended by DALHAP Investments Ltd., Effektiv Spenden, Ergo Impact, Founders Pledge, Charlottes och Fredriks Stiftelse, Good Ventures, Jaan Tallinn, Longview, Open Philanthropy, and Waking Up Foundation. A complete list of donors and funders is available at www.rand.org/TASP. RAND donors and grantors have no influence over research findings or recommendations.

Acknowledgments

I acknowledge Jim Mitre, Eric Drexler, Matan Chorev, Nidhi Kalra, Richard Danzig, Ben Chang, Jeff Alstott, Mike Mazarr, David Frelinger, Mike Lostumbo, Miles Brundage, Andrew Critch, Greg McElvey, Gopal Sarma, Henry Willis, Buzz Phillips, and Edward Geist for conversations that shaped these ideas.

About the Author

Joel B. Predd is a senior engineer at RAND and director of the Geopolitics of AGI Initiative. His current research focuses on the geopolitics of artificial intelligence and on the intersection of the security and economic competition with China. Predd holds a Ph.D. in electrical engineering.

Contents

About This Working Paper iii
Contents v
Figures vi
Riding the Coming Wave: The Urgent Need for Contingency Plans for AGI Futures 1
Introduction and Diagnosis 1
Implications 6
Prescription 8
Riding the Coming Wave 10
References 11

Figures

Figures

Figure 1: Contingency-Signposts as Windows into an AGI Future 9

Riding the Coming Wave: The Urgent Need for Contingency Plans for AGI Futures

Introduction and Diagnosis

The accelerating pace of AI development and diffusion, and the inherently limited governmental capacity for anticipatory policymaking are critical—albeit only implicitly acknowledged—attributes of the strategic context within which strategies to respond to and take advantage of artificial general intelligence (AGI) are being developed and evaluated. As a result, this working paper assumes that a national strategy to adapt to a future with AGI plausibly will emerge from government's responses to events rather than from a top-down analytic or policymaking process. The choices that U.S. government officials and business leaders make in response to such events could influence the future just as much, if not more, than institutions or policies intended to govern the development and diffusion of AGI.

Accelerating Pace of AI Development and Diffusion

Since the dawn of the deep learning revolution with the success of AlexNet in 2012, the pace of AI progress and diffusion has accelerated.[1] Much ink has been spilled predicting timelines for the emergence of AGI. From the perspective of developing strategy and policy for AGI, regardless of the accuracy of those predictions, policymakers plausibly will not enjoy the benefit of time to develop an architecture for secure and stable AI futures the way that the U.S. government did in addressing the development of nuclear weapons during the early Cold War.

In the late 1940s, the United States enjoyed an atomic monopoly; the Soviet Union, despite beating the United States to space with Sputnik, only developed the ability to impose assured destruction on the United States in the mid-1960s with the maturation of its missile force.[2] This provided a period of almost *two decades* during which analysts at RAND and elsewhere developed the concepts and frameworks that the United States adopted to navigate the challenges of the nuclear age. More than a decade passed between when RAND first contemplated intercontinental ballistic missiles (ICBMs) in its inaugural 1946 report *Preliminary Design of an Experimental World-Circling Spaceship* and when the Soviet Union tested the R-7 ICBM in

[1] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097-1105.

[2] "Draft Memorandum from Secretary of Defense McNamara to President Kennedy," November 14, 1963, in David W. Mabon, ed., Foreign Relations of the United States, 1961–1963, Vol. VIII, National Security Policy, U.S. Government Printing Office, 1996.

1957.[3] RAND studies in the early 1950s explored the possibilities of reconnaissance satellites (which became a reality with Corona in 1960) and identified the importance of a secure second-strike force for nuclear deterrence.[4] If today's analysts and policymakers accept that (1) the nuclear era is a fraught but nonetheless illuminating historical analogue and (2) AGI may be imminent, as some experts and leaders of top AI labs predict,[5] then by that analogy serious development of a security architecture for AGI should have begun *prior* to the deep learning revolution—when AGI was not more than a twinkle in most technologists' eyes.[6]

In fact, foundational intellectual work by various scholars on the geopolitics of AI *has been* conducted and published—for example, by the Future of Humanity Institute,[7] among others, who more presciently foresaw the future we confront today. National security researchers began to take more note when the Center for Security and Emerging Technology (CSET) was established in 2019.[8] RAND itself did not launch the Geopolitics of AGI Initiative until December 2023, after many leading AI experts had publicly raised questions about the national security implications of AGI and around the time that trendlines in AI capabilities were accelerating.[9] Very recently, the mainstream national security community has begun a nascent debate over the need for—and components of—a strategy to address a future in which transformative AI has arrived. Prominent national security professionals,[10] technology leaders, [11] and leading AI

---

[3] Clauser, F. H, Rand Corporation, and Project Rand. 1998. Preliminary Design of an Experimental World-Circling Spaceship. Special anniversary edition. Santa Monica, Calif: RAND, pp. 9-10.

[4] Merton E. Davies and William R. Harris, RAND's Role in the Evolution of Balloon and Satellite Observation Systems and Related U.S. Space Technology, RAND Corporation, R-3692, 1988; E. S. Quade, "The Selection and Use of Strategic Air Bases: A Case History," in E. S. Quade, ed., Analysis for Military Decisions, RAND Corporation, R-387-PR, 1964.

[5] See, for example, John Koetsier, "OpenAI CEO Sam Altman: 'We Know How to Build AGI,'" *Forbes,* January 6, 2025. As of May 12, 2025: https://www.forbes.com/sites/johnkoetsier/2025/01/06/openai-ceo-sam-altman-we-know-how-to-build-agi/?utm_source=chatgpt.com; Anthropic, "Paris AI Summit," *Anthropic News*, undated. As of May 12, 2025: https://www.anthropic.com/news/paris-ai-summit?utm; and "AI Has the Potential to Cure All Diseases, Says DeepMind Chief," *The Times*, January 6, 2025. As of May 12, 2025: https://www.thetimes.com/business-money/technology/article/ai-has-the-potential-to-cure-all-diseases-says-deepmind-chief-9hbdp5kpm?utm_source=chatgpt.com&region=global.

[6] The thought experiment is fraught if not misleading—but any other technology analogy will suffer in different ways. Thus, its purpose is simply to juxtapose timelines to highlight implications for strategy.

[7] See, for example, Stuart Armstrong, Nick Bostrom, and Carl Shulman, Racing to the Precipice: Artificial Intelligence Development, Future of Humanity Institute, Technical Report 2013-1, October 2013; Ben Garfinkel and Allan Dafoe, "How Does the Offense-Defense Balance Scale?" Journal of Strategic Studies, Vol. 42, No. 6, 2019; Waqar Zaidi and Allan Dafoe, International Control of Powerful Technology: Lessons from the Baruch Plan for Nuclear Weapons, Future of Humanity Institute, Centre for the Governance of AI, March 2021.

[8] Ben Buchanan, The AI Triad and What It Means for National Security Strategy, Center for Security and Emerging Technology, August 2020.

[9] Leopold Aschenbrenner, "Situational Awareness: The Decade Ahead," June 2024, https://situational-awareness.ai.

[10] Philip Zelikow, Mariano-Florentino Cuéllar, Eric Schmidt, and Jason Matheny, *Defense Against the AI Dark Arts: Threat Assessment and Coalition Defense*, December 2024.

[11] Dan Hendrycks, Eric Schmidt, and Alexandr Wang, *Superintelligence Strategy: Expert Version*, arXiv:2503.05628v1, March 7, 2025, https://www.nationalsecurity.ai.

companies[12] are offering up concepts and plans for a national strategy and plan to implement it. Many of these reports propose useful no-regret options, and others offer bold, if controversial,[13] proposals. Most of these contributions focus on strategies for securely and safely *developing and diffusing* advanced AI rather than living in futures defined by transformative AI.

Alas, the U.S. government still lacks coherent strategies, policies, and essential capabilities to navigate the multidimensional and enormously complex national security problems posed by the future of transformative AI.[14]

Limited Capacity for Anticipatory Policymaking

A related but separate issue is that, even if the seed of a sensible and feasible strategy concept were to germinate within the analytic community, at RAND or elsewhere, the U.S. government is unlikely to have the absorptive capacity for anticipatory policymaking to allow those seeds to grow. I speculate that three issues are in play:[15]

- *Fundamental uncertainty and limited imagination*. While many leading AI experts and researchers take the transformative potential of AI as given (AGI "pilled," as has become the slang *de jure*), there remain myriad resolvable and deep uncertainties. Will AI labs resolve the alignment and control problem? Will there be a decisive first-mover advantage? How quickly will AI diffuse? Will the gap between open and close source

---

[12] Anthropic, "Anthropic's Recommendations to OSTP for the U.S. AI Action Plan," webpage, March 6, 2025, https://www.anthropic.com/news/anthropic-s-recommendations-ostp-u-s-ai-action-plan; OpenAI, "OpenAI's proposals for the U.S. AI Action Plan," webpage, March 12, 2025, https://openai.com/global-affairs/openai-proposals-for-the-us-ai-action-plan/; Kent Walker, "Google's Comments on the U.S. AI Action Plan," *The Keyword* (Google blog), March 13, 2025, https://blog.google/outreach-initiatives/public-policy/google-us-ai-action-plan-comments.

[13] Iskander Rehman, Karl P. Mueller, and Michael J. Mazarr, "Seeking Stability in the Competition for AI Advantage," commentary, RAND Corporation, March 13, 2025, https://www.rand.org/pubs/commentary/2025/03/seeking-stability-in-the-competition-for-ai-advantage.html.

[14] Jim Mitre and Joel B. Predd, *Artificial General Intelligence's Five Hard National Security Problems*, RAND Corporation, PE-A3691-4, February 2025, www.rand.org/t/PEA3691-4.

[15] On this score, it seems worth citing Winston Churchill, "Fifty Years Hence," December 1931, in which Churchill prompts readers to imagine a future in which profound advancements in science and technology challenge social and political realities. For example, he writes, "Mankind has sometimes travelled forwards and sometimes backwards, or has stood still even for hundreds of years. […] What is it that has produced this new prodigious speed of man? Science is the cause. […] It is this power called Science which has laid hold of us, conscripted us into its regiments and batteries, set us to work upon its highways and in its arsenals; rewarded us for our services, healed us when we were wounded, trained us when we were young, pensioned us when we were worn out. None of the generations of men before the last two or three were ever gripped for good or ill and handled like this." He continues: "There are two processes which we adopt consciously or unconsciously when we try to prophesy. We can seek a period in the past whose conditions resemble as closely as possible those of our day, and presume that the sequel to that period will, save for some minor alterations, be repeated. Secondly, we can survey the general course of development in our immediate past, and endeavour to prolong it into the near future. The first is the method of the historian; the second that of the scientist. Only the second is open to us now, and this only in a partial sphere. By observing all that Science has achieved in modern times, and the knowledge and power now in her possession, we can predict with some assurance the inventions and discoveries which will govern our future. We can but guess, peering through a glass darkly, what reactions these discoveries and their applications will produce upon the habits, the outlook and the spirit of men."

models contract or expand? And, perhaps first and foremost, will the scaling paradigm prevail? These uncertainties are not likely to be resolved before consequences of advanced AI for national security manifest. What's more, they are confounded by the absence of a clarifying moment like the splitting of the atom, the Trinity test, a memo from Bruno Augustein,[16] or the Berlin and Cuban Missile crises that provides the evidence that galvanizes the attention and imagination of policymakers and their constituents.

Uncertainty in its most profound form is reflected in recognized ignorance[17] that constrains our imaginations. Jack Clark, a co-founder of Anthropic, observed that the primary constraint on his ability to benefit from an AI tool was his own imagination.[18] Often, our intuitions are incoherent in a sense argued by Eric Drexler, a leading AI expert and researcher:[19] Many of us are guilty of simultaneously presuming that (1) AI will utterly transform economies, militaries, and society yet also that (2) there will be continuity in geopolitics and the instruments of national power that allow the United States to remain competitive globally and to manage that transition to AGI at home and abroad. I hypothesize that the futures many strategists have in mind when discussing the most advanced forms of AI aptly might be titled "2010 with AGI characteristics."

- *Legitimate disagreement on fundamental trade-off*s. AI has potentially enormous upsides[20] and downslides; in fact, it can have both, to different actors and in different ways. Consequently, most policies that govern AI imply some constraint on development or diffusion of the technology in service of security or safety and thus carry some cost for the economic growth or innovation. Similarly, polices that prioritize economic growth or innovation may carry negative implications for job displacement or safety. The means to reaching AGI also often involve trade-offs—for example, efforts to accelerate development through data center build-outs have implications for energy use and privacy, and vice versa. Often, benefits are clear while the risks appear more speculative, but in any case there is legitimate debate about the fundamental trade-offs inherent in most AI strategies and policies. There are objective facts that can be established, but subjective judgments about policy trade-offs are inescapable.[21]

---

[16] Bruno Augenstein, A Revised Development Program for Ballistic Missiles of Intercontinental Range, RAND Corporation, SM-21, 1954, https://www.rand.org/pubs/special_memoranda/SM21.html.

[17] Scott Janzwood, "Confidence Deficits and Reducibility: Toward a Coherent Conceptualization of Uncertainty Level," Risk Analysis, August 21, 2022.

[18] Jack Clark, "Import AI 397: DeepSeek Means AI Proliferation Is Guaranteed; Maritime Wardrones; and More Evidence of LLM Capability Overhangs," Import AI, Substack, January 27, 2025.

[19] Eric Drexler, "Incoherent AI Scenarios Are a Threat," AI Prospects: Toward Global Goal Alignment, Substack, October 24, 2024.

[20] Dario Amodei, Machines of Loving Grace: How AI Could Transform the World for the Better, October 2024.

[21] A related issue is that not only do we lack domestic consensus on AI, but we also lack allied consensus and indeed still are inhibited by differences in national approaches to tech governance and deployment with some of our closest allies and partners. For example, see Krystyna Marcinek, Karlyn D. Stanley, Gregory Smith, Paul Cormarie, Salil Gunashekar, "The EU AI Act Bans Certain AI Practices and Sets Requirements for Limited-Risk Systems," RAND Corporation, RR-A3243-3, 2023. As of May 13, 2025: https://www.rand.org/pubs/research_reports/RRA3243-3.html#:~:text=The%20EU%20AI%20Act% and Hiroki Habuka, "Japan's Approach to AI Regulation and Its Impact on the 2023 G7 Presidency," Center for Strategic and International Studies, February 14, 2023. As of May 13, 2025: https://www.csis.org/analysis/japans-approach-ai-regulation-and-its-impact-2023-g7-presidency.

- *Opportunity costs for finite political, human, physical, and financial capital*. Given fundamental uncertainty and disagreement about trade-offs, most bold actions require policymakers to accept costs in terms of political, human, physical, or financial capital. Those resources are finite, and there are ample opportunities to use those resources for myriad other aims, often with fewer uncertainties, broader consensus, or lesser costs. Investing in a fleet of autonomous, relatively low-cost unmanned aerial vehicles[22] may seem to some like a more worthwhile investment than building a "country of geniuses in a datacenter."[23] What's more, policymakers have limited agency—whereas the U.S. government controlled nuclear technology in the prior era, it now neither owns the AI technology nor the AI models and is limited in its ability to direct use cases,[24] imbuing a degree of passivity and reactionism.

Each of these barriers presents challenges to *prescribing* strategy in a real-world context; there are also sobering challenges to *executing* any strategy. For example, the then-bipartisan CHIPS and Science Act is still in execution four years after the original legislative proposals.[25] Since its inception in 2022, the DoD Chief Digital and AI Office—nominally the DoD point for

---

[22] Thomas Hamilton and David A. Ochmanek, Operating Low-Cost, Reusable, Unmanned Aerial Vehicles in Contested Environments: Preliminary Evaluation of Operational Concepts, RAND Corporation, RR-4407, 2020, www.rand.org/t/RR4407 and Christian Brose, "The New Revolution in Military Affairs," Foreign Affairs, 2023. As of May 13, 2025: https://www.foreignaffairs.com/united-states/new-revolution-military-affairs.

[23] In *Machines of Loving Grace*, Amodei describes powerful AI as an model that is "smarter than a Nobel Prize winner across most relevant fields," "has all the 'interfaces' available to a human working virtually," "can be given tasks that take hours, days, or weeks to complete," "can control existing physical tools," and can do all of this with a scope and scale unmatched by humans or, as he puts it, a "country of geniuses in a datacenter."

[24] Ben Buchanan has emphasized "that this is the first revolutionary technology that is not funded by the Department of Defense, basically. And if you go back historically, over the last hundred years or so, nukes, space, the early days of the internet, the early days of the microprocessor, the early days of large-scale aviation, radar, the global positioning system — the list is very, very long — all of that tech fundamentally comes from Department of Defense money. It's the private sector inventing it, to be sure. But the central government role gave the Department of Defense and the U.S. government an understanding of the technology that, by default, it does not have in AI. It also gave the U.S. government a capacity to shape where that technology goes that, by default, we don't have in AI."

[25] The CHIPS and Science Act of 2022 (Public Law 116-167) was developed in response to growing concerns over U.S. reliance on foreign semiconductor supply chains, particularly from Taiwan and South Korea, and the impact of global chip shortages on industries and national security. Legislative efforts began in 2020 with the introduction of the CHIPS for America Act, which laid the groundwork for semiconductor funding and was later incorporated into broader competitiveness bills like the U.S. Innovation and Competition Act and the America COMPETES Act in early 2022. After months of bipartisan negotiations, Congress finalized the CHIPS and Science Act, passing it in the Senate on July 27, 2022 (64-33 vote) and in the House on July 28, 2022 (243-187 vote), before President Joe Biden signed it into law on August 9, 2022. The act allocated $52.7 billion for semiconductor manufacturing and R&D, including $39 billion in direct incentives, $75 billion in loans, and $13.2 billion for workforce and R&D programs. The Department of Commerce began implementation in late 2022, launching the CHIPS Program Office and opening funding applications for semiconductor companies. Throughout 2023, major companies, such as Intel, TSMC, Samsung, and Micron, announced multibillion-dollar U.S. semiconductor facility investments backed by CHIPS Act funding. By 2024, further initiatives focused on R&D, workforce development, and the launch of the National Semiconductor Technology Center. The act is a cornerstone of U.S. industrial policy, aiming to restore domestic semiconductor leadership, enhance technological innovation, and strengthen economic and national security in the face of global competition.

AI adoption—has been through two directors and multiple reorganizations, not counting the reorganization of its predecessor, the Joint AI Center.[26]

Implications

Nothing New?

One could argue that this diagnosis is neither new nor consequential. First, historical events *typically* crystallize a problem for decisionmakers that theretofore were understood by the analyst class. But for the Cuban Missile Crisis or Berlin Crisis, would there have been a window of opportunity to institute arms control elements of the nuclear security architecture? (How could it happen any other way?) Second, our national response to myriad crises—9/11, the financial crisis of 2008, the COVID-19 pandemic, the 2022 Russian invasion of Ukraine—were possible only because our institutions were skilled in the practice of contingency planning and response. Finally, flexibility, agility, and adaptiveness are well-recognized features of proactive and adaptive strategy.[27] The defense planning apparatus within the Pentagon has demonstrated that, despite its flaws, a modern bureaucracy can plan for the future, even if its inability to simultaneously plan for responses to *any and all* contingencies—such as responses to biothreats, climate change, or a combination of these and more—provide reason for more pessimism.

Indeed, I acknowledge that the separate but related issues of an unknown timeline and uncertainty about the effects of AGI seem navigable on their own. Rapid technological change could be manageable if we had clarity on the challenge, consensus on necessary trade-offs, and means to overcome costs. On the other hand, given sufficient time, there are ample methods for strategic planning that could allow us to cope with uncertainty, forge consensus on trade-offs, and support robust decisionmaking under deep uncertainty.[28]

---

[26] The Department of Defense Chief Digital and AI Officer (CDAO) role was established in December 2021 to centralize DoD's AI, data, and digital transformation efforts by integrating the Joint AI Center (JAIC), Defense Digital Service (DDS), and the Office of Advancing Analytics. The office became fully operational in June 2022. Craig Martell, a former AI executive at Lyft and LinkedIn, was appointed as the first CDAO in April 2022, leading efforts to enhance AI-driven military operations, data governance, and Joint All-Domain Command and Control (JADC2). After his departure in April 2024, he was succeeded by Radha Iyengar Plumb, who initiated a restructuring in late 2024 to improve efficiency ahead of a possible administration transition.

[28] Robert J. Lempert, Steven W. Popper, and Steven C. Bankes, Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis, RAND Corporation, MR-1626-RPC, 2003, www.rand.org/t/MR1626; Paul K. Davis, Analytic Architecture for Capabilities-Based Planning, Mission-System Analysis, and Transformation, RAND Corporation, MR-1513-OSD, 2002, www.rand.org/t/MR1513; Hugh Courtney, Jane Kirkland, and Patrick Viguerie, "Strategy Under Uncertainty," Harvard Business Review, November–December 1997.

The Importance of Contingency

But taken together, these two forces—the accelerating technological progress in frontier AI and limited capacity for anticipatory policymaking—are defining attributes of the strategic environment. First, they suggest that the U.S. strategy, policies, and institutions needed for an AGI future plausibly will not emerge from a top-down analytical and policy process on timelines that matter. Second, the events subject to contingency planning may be harbingers of a coming wave[29]—not merely disruptions to the status quo—that require bold moves that would likely be too risky, costly, or infeasible to implement prior to the event.

Under these conditions, the choices the U.S. government makes in response to such events could influence the future just as much, if not more, than institutions or policies that are designed to govern the development and diffusion of AGI. Thus, AI experts and policymakers should be prepared to respond to events—events that resolve uncertainty with evidence, provide a basis for overcoming disagreement, and galvanize imaginations to allocate requisite resources and attention—with plans that may appear infeasible or unwise amid today's uncertainties, debates, or trade-offs but which seize the opportunity presented by unfolding events to shape AI futures to ends desirable for the United States and for humanity.[30] Given the timelines, these plans and capabilities most plausibly must be developed beforehand.

For example, experts are concerned about the prospect of AI providing a first-mover military advantage, conjuring scenarios in which an actor secures a decisive new capability, say, in offensive or defensive cyber operations that suddenly upsets the military balance or undermines deterrence.[31] This leads to questions about the feasibility and durability of such a prospective first-mover advantage, and considerations about strategies to secure such an advantage and the risks of being late. But what will be done if that moment arrives before the U.S. government is ready? What are governments' plans to deter potential cyberattacks supercharged by transformative AI, and what capabilities would be needed to establish a robust deterrent? How should the U.S. government respond if the United States is attacked? What should the U.S. government communicate—or not communicate—to adversaries about its own willingness to use AI in this manner?

In another example, experts are concerned about scenarios in which AI developers lose control of an agentic AI.[32] This leads one to questions about how we can align or control AI in

---

[29] Mustafa Suleyman, with Michael Bhaskar, The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma, Crown, 2003.

[30] Eric Drexler, "AI-Driven Strategic Transformation: Preparing to Pivot," AI Prospects: Toward Global Goal Alignment, Substack, December 9, 2024.

[31] Jim Mitre and Joel B. Predd, *Artificial General Intelligence's Five Hard National Security Problems,* RAND Corporation, PE-A3691-4, 2025. As of May 12, 2025: https://www.rand.org/pubs/perspectives/PEA3691-4.html

[32] For example, see Jonathan Zittrain, "The Safety Risks of AI Agents," *The Atlantic,* July 2024. As of May 12, 2025: https://www.theatlantic.com/technology/archive/2024/07/ai-agents-safety-risks/678864/ or Yoshua Bengio, "How Rogue AIs May Arise," *yoshuabengio.org*, May 22, 2023. As of May 12, 2025: https://yoshuabengio.org/2023/05/22/how-rogue-ais-may-arise/

advance and, more broadly, govern AI development and diffusion with security and safety in mind. But what will be done if that moment arrives before we are ready? For example, what would the U.S. government's plan be to contain the outbreak, reset the global digital infrastructure to a known state, and negotiate an international agreement on global AI governance, after having received evidence that resolves, beyond all doubt, that humans can, indeed, create autonomous, goal-seeking AI?

Policymakers need options to respond to—not just to prevent or anticipate—acute threats to national security and moments when their choices are likely to determine our future and probably need to be made rapidly. What we are imagining is akin to having a plan for a comprehensive global counterterrorism campaign *in advance* of the catastrophic terrorist attacks of 9/11—a plan that considered the necessary military doctrine, technologies, intelligence requirements, institutions, and legal authorities, and their second- and third-order strategic effects.

Prescription

What would it mean to embrace this diagnosis in our strategic concepts—rather than merely heed it as a call for strategists to get their act together?

First, embracing this diagnosis would mean developing and proactively *identifying contingency-signposts*—events that both represent threats to the status quo *and* the coming wave. In other words, events that possess acute symptoms for the economy, national security, or society that must be addressed directly and urgently. But those also represent signposts that uncertainty has resolved to reveal clues about the future that may soon emerge.[33] There are many examples of such contingency-signposts:

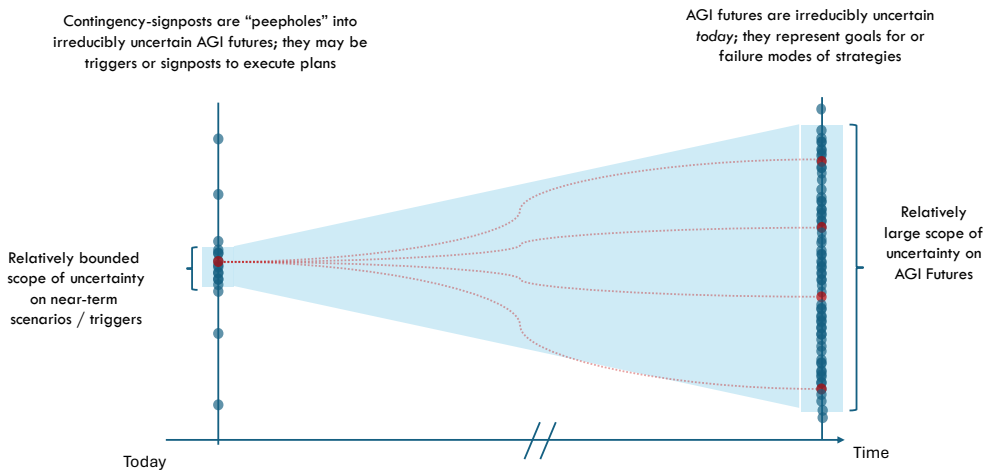
- a Chinese Communist Party announcement of a national plan to sprint to Artificial Super Intelligence (ASI);
- demonstration by a leading AI company that it has crossed a performance threshold claimed to mark the development of AGI;
- a sudden uptick in unemployment due to the adoption of AI in particular sectors or job class;
- revelation of a more lethal variant of influenza engineered by a nonstate actor with help from AI;
- demonstration of new intelligence fusion capability that exposes the location of stealthy delivery platforms;
- digital AI companions exerting agency and expanding to "infest" global digital infrastructure on which financial, energy, telecommunications, and health care systems depend; or

[33] The contingency signposts that we envision may be neither fully fleshed out scenarios nor complete descriptions of the future; events that may motivate a response may be rife with their own forms of uncertainty.

- adversary development of an AI-enabled cyber capability that is used to harden previously vulnerable targets and to hold at risk vulnerable critical infrastructure.

Each of these may represent clarifying moment, like the splitting of the atom, the Trinity test, or the Berlin and Cuban Missile Crisis. Figure 1 illustrates the concept of contingency-signpost in graphic form.

Figure 1: Contingency-Signposts as Windows into an AGI Future



Second, embracing this diagnosis would require *developing strategy frameworks and "break glass" plans with multiple levels of ambition and resolution*. Given the scenario, the first step is to develop a strategy framework that articulates (1) the near- and long-term goals and priorities prompted by the event and (2) the key choices that policymakers must confront. This is necessary to help policymakers see through the details of the scenario to identify the central choices they must make. Given such a strategy framework, planning should address two levels. The first level of planning must treat the acute issue—like a patient who enters an emergency room bleeding. The second level of planning should embrace the moment of clarity to put in place policies, institutions, and capabilities that heretofore were infeasible or unwise amid the uncertainties, debates, or opportunity costs that preceded—like a patient who has come to terms with a needed lifestyle change.

Finally, embracing this diagnosis requires *developing and exercising capabilities and competitiveness needed to execute "break glass" plans.* Some of the plans may requires capabilities (systems, organizations, authorities, etc.) that do not yet exist, either to prevent an event from occurring in the first place, to improve information available to decisionmakers choosing to execute elements of a plan, or to execute the plan itself. The final step is to develop those capabilities in advance and to practice the competencies they require.

Riding the Coming Wave

To be certain, the United States—its government, private sector, and civil society—cannot give up on strategy development. At RAND, the Geopolitics of AGI Initiative is developing and evaluating potential long-term strategies to safeguard and advance U.S. interests on the path to and in a world with advanced AI. These strategies will identify key assumptions, strategic objectives, foundational questions and trade-offs, and core elements of a policy approach.

But policymakers cannot assume that the strategies, institutions, and polices will be fully developed, much less instituted, in time, or that these strategies will prove the paper they are written on. Thus, the Initiative is also identifying contingency-signposts; developing concrete, plausible scenarios; and convening weekly National Security Council simulations with RAND researchers, external experts, and policymakers to stress-test and sharpen potential courses of action; help inform plans, capabilities, and recommended near-term actions to enhance readiness; and deepen knowledge and familiarity with the key issues between key organizations and experts.

Above all, this diagnosis reflects the reality that time may be the most valuable resource in the current moment. Developing plans to implement strategy embraces a strategic logic of buying time.[34] As Eisenhower said, "Plans are worthless, but planning is everything."

[34] Kathryn E. Bouskill, Seifu Chonde, and William Welser IV, Speed and Security: Promises, Perils, and Paradoxes of Accelerating Everything, RAND Corporation, PE-274-RC, May 2018, www.rand.org/t/PE274.

References

"AI Has the Potential to Cure All Diseases, Says DeepMind Chief," *The Times*, January 6, 2025. As of May 12, 2025: https://www.thetimes.com/business-money/technology/article/ai-has-the-potential-to-cure-all-diseases-says-deepmind-chief-9hbdp5kpm?utm_source=chatgpt.com&region=global

Amodei, Dario, Machines of Loving Grace: How AI Could Transform the World for the Better, October 2024.

Anthropic, "Paris AI Summit," *Anthropic News*, undated. As of May 12, 2025: https://www.anthropic.com/news/paris-ai-summit?utm

Anthropic, "Anthropic's Recommendations to OSTP for the U.S. AI Action Plan," webpage, March 6, 2025, https://www.anthropic.com/news/anthropic-s-recommendations-ostp-u-s-ai-action-plan

Armstrong, Stuart, Nick Bostrom, and Carl Shulman, Racing to the Precipice: Artificial Intelligence Development, Future of Humanity Institute, Technical Report 2013-1, October 2013

Aschenbrenner, Leopold, "Situational Awareness: The Decade Ahead," June 2024, https://situational-awareness.ai.

Augenstein, Bruno, A Revised Development Program for Ballistic Missiles of Intercontinental Range, RAND Corporation, SM-21, 1954, https://www.rand.org/pubs/special_memoranda/SM21.html.

Bengio, Yoshua, "How Rogue AIs May Arise," *yoshuabengio.org*, May 22, 2023. As of May 12, 2025: https://yoshuabengio.org/2023/05/22/how-rogue-ais-may-arise/

Bouskill, Kathryn E., Seifu Chonde, and William Welser IV, Speed and Security: Promises, Perils, and Paradoxes of Accelerating Everything, RAND Corporation, PE-274-RC, May 2018, www.rand.org/t/PE274

Brose, Christian, "The New Revolution in Military Affairs," Foreign Affairs, 2023. As of May 13, 2025: https://www.foreignaffairs.com/united-states/new-revolution-military-affairs

Buchanan, Ben, The AI Triad and What It Means for National Security Strategy, Center for Security and Emerging Technology, August 2020

Churchill, Winston, "Fifty Years Hence," December 1931.

Clark, Jack, “Import AI 397: DeepSeek Means AI Proliferation Is Guaranteed; Maritime Wardrones; and More Evidence of LLM Capability Overhangs,” Import AI, Substack, January 27, 2025.

Clauser, F. H, Rand Corporation, and Project Rand. 1998. Preliminary Design of an Experimental World-Circling Spaceship. Special anniversary edition. Santa Monica, Calif: RAND, pp. 9-10.

Courtney, Hugh, Jane Kirkland, and Patrick Viguerie, “Strategy Under Uncertainty,” Harvard Business Review, November–December 1997

Davies, Merton E., and William R. Harris, RAND’s Role in the Evolution of Balloon and Satellite Observation Systems and Related U.S. Space Technology, RAND Corporation, R-3692, 1988; E. S. Quade, “The Selection and Use of Strategic Air Bases: A Case History,” in E. S. Quade, ed., Analysis for Military Decisions, RAND Corporation, R-387-PR, 1964.

Davis, Paul K., Analytic Architecture for Capabilities-Based Planning, Mission-System Analysis, and Transformation, RAND Corporation, MR-1513-OSD, 2002, www.rand.org/t/MR1513

“Draft Memorandum from Secretary of Defense McNamara to President Kennedy,” November 14, 1963, in David W. Mabon, ed., Foreign Relations of the United States, 1961–1963, Vol. VIII, National Security Policy, U.S. Government Printing Office, 1996.

Drexler, Eric, ”Incoherent AI Scenarios Are a Threat,” AI Prospects: Toward Global Goal Alignment, Substack, October 24, 2024.

Drexler, Eric, “AI-Driven Strategic Transformation: Preparing to Pivot,” AI Prospects: Toward Global Goal Alignment, Substack, December 9, 2024.

Garfinkel, Ben, and Allan Dafoe, “How Does the Offense-Defense Balance Scale?” Journal of Strategic Studies, Vol. 42, No. 6, 2019

Hamilton, Thomas, and David A. Ochmanek, Operating Low-Cost, Reusable, Unmanned Aerial Vehicles in Contested Environments: Preliminary Evaluation of Operational Concepts, RAND Corporation, RR-4407, 2020, www.rand.org/t/RR4407

Hendrycks, Dan, Eric Schmidt, and Alexandr Wang, *Superintelligence Strategy: Expert Version*, arXiv:2503.05628v1, March 7, 2025, https://www.nationalsecurity.ai.

Hiroki Habuka, "Japan's Approach to AI Regulation and Its Impact on the 2023 G7 Presidency," Center for Strategic and International Studies, February 14, 2023. As of May 13, 2025: https://www.csis.org/analysis/japans-approach-ai-regulation-and-its-impact-2023-g7-presidency

Janzwood, Scott, "Confidence Deficits and Reducibility: Toward a Coherent Conceptualization of Uncertainty Level," Risk Analysis, August 21, 2022.

Koetsier, John, "OpenAI CEO Sam Altman: 'We Know How to Build AGI,'" *Forbes,* January 6, 2025. As of May 12, 2025: https://www.forbes.com/sites/johnkoetsier/2025/01/06/openai-ceo-sam-altman-we-know-how-to-build-agi/?utm_source=chatgpt.com

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems,* 25, 1097-1105.

Lempert, Robert J., Steven W. Popper, and Steven C. Bankes, Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis, RAND Corporation, MR-1626-RPC, 2003, www.rand.org/t/MR1626

Marcinek, Krystyna, Karlyn D. Stanley, Gregory Smith, Paul Cormarie, Salil Gunashekar, "The EU AI Act Bans Certain AI Practices and Sets Requirements for Limited-Risk Systems," RAND Corporation, RR-A3243-3, 2023. As of May 13, 2025: https://www.rand.org/pubs/research_reports/RRA3243-3.html#:~:text=The%20EU%20AI%20Act%

Mitre, Jim, and Joel B. Predd, *Artificial General Intelligence's Five Hard National Security Problems,* RAND Corporation, PE-A3691-4, 2025. As of May 12, 2025: https://www.rand.org/pubs/perspectives/PEA3691-4.html

OpenAI, "OpenAI's proposals for the U.S. AI Action Plan," webpage, March 12, 2025, https://openai.com/global-affairs/openai-proposals-for-the-us-ai-action-plan/

Rehman, Iskander, Karl P. Mueller, and Michael J. Mazarr, "Seeking Stability in the Competition for AI Advantage," commentary, RAND Corporation, March 13, 2025, https://www.rand.org/pubs/commentary/2025/03/seeking-stability-in-the-competition-for-ai-advantage.html

Suleyman, Mustafa, with Michael Bhaskar, The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma, Crown, 2003

Walker, Kent, "Google's Comments on the U.S. AI Action Plan," *The Keyword* (Google blog), March 13, 2025, https://blog.google/outreach-initiatives/public-policy/google-us-ai-action-plan-comments

Zaidi, Waqar and Allan Dafoe, International Control of Powerful Technology: Lessons from the Baruch Plan for Nuclear Weapons, Future of Humanity Institute, Centre for the Governance of AI, March 2021.

Zelikow, Philip, Mariano-Florentino Cuéllar, Eric Schmidt, and Jason Matheny, *Defense Against the AI Dark Arts: Threat Assessment and Coalition Defense*, December 2024.

Zittrain, Jonathan, “The Safety Risks of AI Agents,” *The Atlantic,* July 2024. As of May 12, 2025: https://www.theatlantic.com/technology/archive/2024/07/ai-agents-safety-risks/678864/