Testimony

JIM MITRE

Artificial General Intelligence's Five Hard National Security Problems

CT-A3914-1
Testimony presented before the U.S. Senate Committee on Armed Services, Cybersecurity Subcommittee on March 25, 2025

For more information on this publication, visit **www.rand.org/t/CTA3914-1**.

Testimonies

RAND testimonies record testimony presented or submitted by RAND associates to federal, state, or local legislative committees; government-appointed commissions and panels; and private review and oversight bodies.

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, please visit www.rand.org/pubs/permissions.

Artificial General Intelligence's Five Hard National Security Problems

Testimony of Jim Mitre[1]
RAND[2]

Before the Committee on Armed Services
Subcommittee on Cybersecurity
United States Senate

March 25, 2025

Chairman Rounds, Ranking Member Rosen, and distinguished members of the committee, thank you for the opportunity to testify today on the national security implications posed by the potential emergence of advanced artificial intelligence (AI) or artificial general intelligence (AGI).[3]

Leading AI labs in the United States, China, and the rest of the world are in hot pursuit of AGI, which would possess human-level or superhuman-level intelligence across a wide variety of cognitive tasks. The pace and potential progress of AGI's emergence—as well as the composition of a post-AGI future—are uncertain and hotly debated.[4] Yet the emergence of AGI is plausible, and the consequences so profound, that the U.S. national security community should take it seriously—and plan for it.

Consider the following: What would the U.S. government do if, in the next few years, a leading AI lab announced that its forthcoming model had the ability to produce the equivalent of

---

[1] The opinions and conclusions expressed in this testimony are the author's alone and should not be interpreted as representing those of RAND or any of the sponsors of its research.

[2] 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. RAND's mission is enabled through its core values of quality and objectivity and its commitment to integrity and ethical behavior. RAND subjects its research publications to a robust and exacting quality-assurance process; avoids financial and other conflicts of interest through staff training, project screening, and a policy of mandatory disclosure; and pursues transparency through the open publication of research findings and recommendations, disclosure of the source of funding of published research, and policies to ensure intellectual independence. This testimony is not a research publication, but witnesses affiliated with RAND routinely draw on relevant research conducted in the organization.

[3] This testimony is drawn from a paper on the topic I drafted with my colleague Joel B. Predd. See Jim Mitre and Joel B. Predd, *Artificial General Intelligence's Five Hard National Security Problems*, RAND Corporation, PE-A3691-4, February 2025, https://www.rand.org/pubs/perspectives/PEA3691-4.html.

[4] Matteo Wong, "The AI Boom Has an Expiration Date," *The Atlantic*, October 17, 2024.

1 million computer programmers as capable as the top 1 percent of human programmers at the touch of a button? The national security implications are profound and could cause a significant disruption of the current cyber offense-defense balance.

At RAND, we are planning for it. Our work has revealed that AGI presents five hard national security problems.

1. Wonder Weapons

First, AGI might enable a significant first-mover advantage via the sudden emergence of a decisive wonder weapon: for example, a capability so proficient at identifying and exploiting vulnerabilities in enemy cyberdefenses that it provides what might be called a *splendid first cyber strike* that completely disables a retaliatory cyberstrike. Such a first-mover advantage could disrupt the military balance of power in key theaters, create a host of proliferation risks, and accelerate technological race dynamics.

2. Systemic Shifts

Second, AGI might cause a systemic shift in the instruments of national power that alters the balance of global power. The history of military innovation suggests that being able to adopt a new technology is more consequential than being the first to achieve a scientific or technological breakthrough.[5] As the U.S., allied, and rival militaries establish access to AGI and adopt it at scale, it could upend military balances by affecting key building blocks of military competition, such as hiders versus finders, precision versus mass, or centralized versus decentralized command and control. States that are better postured to capitalize on—and manage—systemic shifts caused by AGI could have greatly expanded influence.

3. Empowered Nonexperts

Third, AGI might serve as a "malicious mentor" that explains and contextualizes the specific steps that nonexperts can take to develop dangerous weapons, such as virulent cyber malware, widening the pool of people capable of creating such threats. Knowing how to build a weapon of mass destruction is, of course, not the same as actually building it. But technological developments in related fields are lowering execution barriers. For example, advances in agentic AI may assist in performing certain tasks to directly aid a malicious actor in their goals.

4. Artificial Entities

Fourth, AGI might achieve enough autonomy and behave with enough agency to be considered an independent actor on the global stage. Consider an AGI with advanced computer programming abilities that is able to "break out of the box" and engage with the world across

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

cyberspace. It could possess agency beyond human control, operate autonomously, and make decisions with far-reaching consequences. Such an AGI could be misaligned—that is, operate in ways that are inconsistent with the intentions of its human designers or operators, causing unintentional harm. In the extreme, a "loss-of-control" scenario could result, wherein an AGI's pursuit of its desired objectives incentivizes the machine to resist being turned off, counter to human efforts.[6]

5. Instability

Fifth, the pursuit of AGI could foster a period of instability as nations and corporations race to achieve dominance in this transformative technology. This competition might lead to heightened tensions, reminiscent of the nuclear arms race, such that the quest for superiority risks precipitating, rather than deterring, conflict. Nations' perceptions of AGI's feasibility and potential to confer a first-mover advantage could become as critical as the technology itself. The risk threshold for action will hinge not only on actual capabilities but also on perceived capabilities and the intentions of rivals. Misinterpretations or miscalculations, much like those feared during the Cold War, could precipitate preemptive strategies or arms buildups that destabilize global security.

As the U.S. Department of Defense embarks on developing the National Defense Strategy, it will have to grapple with how advanced AI will affect cyber, along with all other domains. The five hard problems that AGI presents to national security can serve as a rubric to evaluate how the strategy addresses the potential emergence of AGI.

Thank you again for the opportunity to testify. I welcome your questions.

[6] Yoshua Bengio, testimony before the U.S. Senate Judiciary Subcommittee on Privacy, Technology, and the Law, July 5, 2023.