Testimony

JEFF ALSTOTT

Preparing the Federal Response to Advanced Technologies

CT-A2953-1
Testimony presented before the U.S. Senate Committee on Homeland Security and Governmental Affairs, Subcommittee on Emerging Threats and Spending Oversight, on September 19, 2023

For more information on this publication, visit www.rand.org/t/CTA2953-1.

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.

Limited Print and Electronic Distribution Rights

Preparing the Federal Response to Advanced Technologies

Testimony of Jeff Alstott[1]
The RAND Corporation[2]

Before the Committee on Homeland Security and Governmental Affairs
Subcommittee on Emerging Threats and Spending Oversight
United States Senate

September 19, 2023

Chair Hassan, Ranking Member Romney, and members of the subcommittee: Good afternoon, and thank you for the opportunity to testify today. I am a senior information scientist with the RAND Corporation, a nonprofit and nonpartisan research organization. Before RAND, I served at the White House as assistant director for technology competition and risks at the Office of Science and Technology Policy and as director for technology and national security at the National Security Council. I also spent time in the intelligence community as a program manager at the Intelligence Advanced Research Projects Activity, with a portfolio that included artificial intelligence (AI), analytic methods, biosecurity, and science and technology forecasting.

For the past 75 years, RAND has conducted research in support of U.S. national security and domestic policy. We manage four federally funded research and development centers for the government focused on national and homeland security. Today, I will focus my comments on

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

[2] The RAND Corporation 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.

how the federal government can respond to emerging threats to national security and public safety posed by broadly capable AI systems, including how they intersect with biosecurity.[3]

Progress in AI has advanced rapidly in recent years, leading to expanded debate among experts about its potential risks. Although AI has the potential to transform entire industries, it could also pose novel threats to national defense and homeland security. AI developers are racing to build increasingly advanced systems, and the drivers of AI progress—including more-efficient algorithms, more-efficient hardware, a better trained and more capable workforce, and greater investment—continue to increase exponentially. Despite this rapid progress, the sciences of interpreting and explaining AI behavior, assessing powerful AI for dangerous capabilities, and designing appropriate guardrails to mitigate harms are all efforts that are still in their infancy. Existing safeguards are still imperfect, and AI models released by leading U.S. companies today can and do still exhibit unsafe and unanticipated behaviors long after they are trained and released. Unless society puts in effective guardrails, broadly capable AI systems could hasten the design and proliferation of bioweapons, cyberweapons, nuclear weapons, progressively more general intelligence, and other threats not yet conceived. If such systems proliferate, it will be very difficult to put the genie back in the bottle, potentially causing irreversible damage.

One particular area of concern is the relationship of advanced AI development with biosecurity. Existing AI models are already capable of assisting nonstate actors with biological attacks that would cause pandemics, including the conception, design, and implementation of such attacks. Without safeguards, the development of ever-more-advanced AI systems will bring ever-greater reductions to the barriers to launch such attacks, until we are at the point in which a lone actor can cause a pandemic, killing millions. This change is occurring at the same time that gene synthesis machines are decreasing in cost, improving in quality and reliability, and proliferating more widely, increasing the number of actors who have the necessary access and ability to create and release new diseases.

Effective oversight of increasingly powerful AI and its potential threats will require visibility into the full AI development lifecycle. This lifecycle begins with large concentrations of AI hardware, with thousands of advanced chips performing a training run costing millions or soon billions of dollars. Once the AI is fully trained, it is made available to the public through a controlled internet interface or by being published online in its entirety, at which point proliferation essentially cannot be stopped. Oversight of each of these stages—AI hardware, training, and release—will be necessary to ensure our national security. These efforts will not come at the cost of U.S. innovation but will bolster U.S. competitiveness by ensuring the safety and reliability of leading U.S. AI products and establishing the United States as a responsible market leader. In addition, domestic oversight, although essential, will not be sufficient alone. We must cooperate with our allies and partners—and communicate responsibly with our competitors—to ensure the safe development of these technologies at the global level.

I will highlight six actions that the federal government could take to mitigate these threats:

---

[3] This testimony builds on previous testimony provided to Congress by RAND's president and chief executive officer. See, for example, Jason Matheny, *Advancing Trustworthy Artificial Intelligence,* RAND Corporation, CT-A2824-1, 2023, https://www.rand.org/pubs/testimonies/CTA2824-1.html.

1. Require that large computing clusters that could be used to train powerful AIs (e.g., high-performance computers with >10,000 advanced AI chips) be reported to the government, have adequate cybersecurity, and have know-your-customer processes for anyone doing a very large computation on them that may be a training run for a powerful AI.
2. Require those making powerful AIs to maintain responsible security procedures during and after the training process to prevent U.S.-made models from being stolen or leaked. The threshold for this requirement could be frontier models trained with >10^26 operations, several times larger than any AI system made before, and should include both those handling the code and those handling the hardware infrastructure.
3. Ensure that these frontier AI development efforts also undergo an independent assessment to determine whether the AI or its proliferation would be a threat to national security, similar to how rocket launches are reviewed by the Federal Aviation Administration. This should include risk assessments prior to model training, safety evaluations and red team tests at regular intervals throughout the training run, and rigorous safety reviews prior to model deployment. Models that are determined to be insufficiently safe could be held for further development and release until safety and security issues are adequately resolved. Conducting safety evaluations in each major stage of the AI development process would help companies detect safety problems early on, when issues are less costly to fix, reducing security risks while saving U.S. companies time and money.
4. Create a safe harbor information-sharing environment for both the private and public sectors to share safety and security problems from their AIs as they identify them and then create solutions.
5. Establish know-your-customer requirements for the providers of gene synthesis services (including cloud laboratory services) and gene synthesis devices (including benchtop synthesizers) to reduce growing biosecurity threats.
6. Require that genetic material synthesized over a threshold (e.g., fragments of >50 base pairs) be screened for pathogenic potential. This should include supporting the development and adoption of a universal, secure, and continuously updated gene synthesis screening mechanism, which would reduce urgent biosecurity threats while decreasing costs for U.S. companies and maintaining U.S. competitiveness in the global bioeconomy.

I thank the subcommittee for the opportunity to testify, and I look forward to answering your questions.