OCTOBER 2020

U.S. Military Investments in Autonomy and AI

Costs, Benefits, and Strategic Effects

CSET Policy Brief

AUTHORS

Margarita Konaev
Husanjot Chahal
Ryan Fedasiuk
Tina Huang
Ilya Rahkovsky

The Department of Defense has an ambitious vision for artificial intelligence. And while most federal agencies have seen their research and development funding decline in 2020, DOD's R&D budget has increased.[1] If leveraged correctly, today's investments in defense research will provide the U.S. military with the AI-enabled capabilities needed to deter adversaries from aggression, fight and win the wars of the future, and cooperate effectively with allies. But where exactly is this investment going? And what benefits and risks might result from developing and fielding autonomous and AI-enabled weapons and systems?

As the strategic competition with China intensifies and allies' defense budgets face mounting pressure, continued U.S. military investment in emerging technologies is more critical than ever. Policymakers need information about DOD's investments in AI to conduct proper oversight and ensure these research efforts support broader strategic goals. Moreover, China, Russia, and other competitors are also investing heavily in military applications of AI. Balancing between the urgency of remaining ahead and building safe and secure AI systems will only become increasingly difficult.

As the U.S. defense community implements its vision for AI, CSET offers a two-part analysis assessing the scope and implications of U.S. military investments in autonomy and AI, focusing on three interconnected elements that form our analytical framework:

- The *technology* element addresses DOD research and development efforts in autonomy and AI;
- The *military capabilities* element speaks to the speed, precision, coordination, reach, persistence, lethality, and endurance enabled by advances in autonomy and AI;
- The *strategic effects* element analyzes how these technological developments and capability enhancements may affect key strategic issues—specifically, deterrence, military effectiveness, and interoperability with allies.

The first report, "U.S. Military Investments in Autonomy and AI: A Budgetary Assessment," centers on the technology element, while the second, "U.S. Military Investments in Autonomy and AI: A Strategic Assessment," covers the military capabilities and strategic effects portions. Both reports draw on publicly available data from the FY2020 research, development, testing, and evaluation (RDT&E) budget justification books of the Army, Air Force, Navy, Marines, and DARPA. We focus specifically on basic, applied, and

advanced research—known jointly as the Science and Technology (S&T) program, which supports the development of new technologies imperative to U.S. military superiority. This analysis is supplemented by an extensive review of strategic and operational literature and scientific research on autonomy and AI.

The following is a summary of both reports, including our assessment of current DOD research investment priorities, trends, and gaps with corresponding recommendations, as well as recommendations for ensuring U.S. military leadership in AI in the short term and the long term.[2]

Current DOD Research Investments: Trends and Gaps

The U.S. military has a wide range of research programs using autonomy and AI in unmanned vehicles and systems, information processing, decision support, targeting functions, and other areas. Yet there are gaps in research on AI not related to autonomy and in investments in basic AI research.

Our results show that estimates of research investments vary depending on definitions and measures. Across the different measurements, however, the data suggests that:

- While most AI-related research efforts are related to autonomy, the majority of autonomy research programs are not related to AI. As such, AI research unrelated to autonomy—and especially autonomy in unmanned systems—receives a relatively small share of the S&T funds directed toward autonomy and AI research.

- Investments in basic AI research are also likely smaller than initially estimated.

The ambiguity about the nature and scope of U.S. military investments in autonomy and AI research makes it difficult to ensure oversight. Moreover, the current U.S. military research on AI may not be sufficiently innovative to fuel the scientific breakthroughs needed to ensure long-term advantage. We therefore offer the following policy recommendations:

- DOD should provide greater clarity about overall funding levels for autonomy and AI, overlap between funding allocated to autonomy research and AI research, and funding for AI-related basic research.

- DOD should leverage its relationships with university-affiliated research centers and national labs to map the landscape of non-

autonomy-related AI and potential military applications, and identify opportunities for additional investment.

Effective human-machine collaboration is key to harnessing the full promise of AI. But gaps in our understanding of trust in human-machine teams can impede progress.

The U.S. military sees many benefits to pairing humans with intelligent technologies and our analysis finds that human-machine collaboration is a crosscutting theme across the different autonomy and AI research programs. The following issues therefore merit attention:

- Trust is essential to human-machine collaboration. Yet in our assessment, few autonomy and AI-related research initiatives reference both trust and human-machine collaboration.

- Gaps in research on the role of trust in human-machine teams can negate the advantages in speed, coordination, and endurance promised by autonomy and AI. This, in turn, could impede U.S. ability to use AI-enabled systems to deter adversaries from aggression, operate effectively on future battlefields, and ensure interoperability with allies.

To safely and effectively employ machines as trusted partners to human operators, the following steps may be necessary:

- DOD should increase investment in multidisciplinary research on the drivers of trust in human-machine teams, specifically under operational conditions.

- DOD should assess the advantages of making trust a consistent theme across autonomy and AI research programs pertaining to human-machine collaboration.

- U.S.-based researchers should collaborate with defense research communities in allied countries on joint research initiatives that assess how cross-cultural variation in trust in human-machine teams may impact interoperability.

AI in the short term

Maximizing advantages: AI applications that enhance military endurance contribute to military effectiveness and readiness, as well as interoperability with allies.

Today's strategic and operational realities put a premium on both operational readiness—supported by elements such as personnel, equipment, supply/maintenance, and training—and endurance—the ability to withstand hostile actions and adverse environmental conditions long enough to achieve mission objectives.[3] There are great opportunities in leveraging existing and relatively safe technologies for logistics and sustainment to streamline personnel management and enhance the functionality and longevity of military equipment. But their potential value is understated.

- AI applications for logistics and sustainment are more than cost saving measures boosting back-office efficiency; they enable military readiness and effectiveness in combat.

- Gaps in endurance capabilities can impair interoperability in multinational coalitions like NATO and undermine the long-term health of U.S. alliances.

Therefore, we offer the following policy recommendations:

- DOD, with coordination support from the Joint Artificial Intelligence Center (JAIC), should calibrate investment in AI applications for endurance as an enabler of military readiness.

- The United States should work closely with allies on AI applications in logistics and sustainment, including joint research and development programs and support for multinational public-private sector partnerships.

Minimizing risks: In the short term, the national security risks of AI have less to do with AI replacing humans and more to do with failure to deliver on technical expectations and with warfighters inexperienced with AI misusing it.

Many of the current U.S. military autonomy and AI R&D projects will never reach fruition; others will fail to scale or be fielded yet rarely used. There are significant technological, organizational, and budgetary barriers to innovation and adoption of new military technologies. As such, a healthy degree of skepticism and tolerance for technical failure are needed.

At the same time, there is also the risk of over-eager adoption. This could result in premature use of AI systems that cannot grasp context or make strategically intricate judgments by warfighters who may not fully understand the potential failure modes of these systems, possibly leading to inadvertent escalation. For instance, employing increasingly autonomous unmanned surface vehicles for military deception, while technically possible, could be destabilizing in highly militarized areas such as the South China Sea. As a result, we recommend the following:

- Security and technology researchers, particularly those affiliated with or advising DOD, should be more explicit about the uncertain pace of progress in specific areas related to autonomy and AI technologies (e.g., autonomous ground combat vehicles and unmanned undersea vehicles).

- The same researchers should differentiate between two types of risks: short-term risks associated with the use and misuse of technologies already in the pipeline, and long-term risks arising from more advanced technologies, which are likely to face development and fielding barriers.

- DOD should coordinate with federally funded research and development centers and university-affiliated research centers to conduct risk assessments and wargames focused explicitly on near-term AI technologies and risks from human-machine interactions.

- JAIC should coordinate with the services and relevant operational commanders to provide thorough training to units operating autonomous and/or AI-enabled systems on the potential failure modes and corresponding risks.

AI in the long term

Robust, resilient, trustworthy, and secure AI systems are key to ensuring long-term military, technological, and strategic advantages.

Numerous S&T programs focus explicitly on strengthening AI robustness and resilience, fortifying security in the face of deceptive and adversarial attacks, and developing systems to behave reasonably and reliably in operational settings (DARPA's "AI Next Campaign" is a prominent example). Yet in our assessment, most S&T programs on autonomy and AI don't mention safety and security attributes. Failure to advance reliable, trustworthy, and resilient

AI systems could adversely affect deterrence, military effectiveness, and interoperability:

- **Deterrence:** The unpredictability and opaqueness of current AI technologies amplify risks of unintended escalation due to the speed of autonomous systems, as well as miscalculations and misperceptions surrounding their potential use.

- **Military Effectiveness:** AI applications that improve situational awareness, decision-making, and targeting processes can enhance precision capabilities and operational effectiveness. But the need for new data input and validation of AI systems deployed in uncertain operational environments raises concerns about unpredictable behavior.

- **Interoperability:** Safety, security, and privacy concerns could stall progress toward AI-enabled coordination between the United States and its allies and undermine the effectiveness of coalitions like NATO.

Based on these findings, we propose the following policy recommendations:

- DOD RDT&E programs should emphasize safety and security across all stages of the AI system lifecycle—from initial design to data/model building, verification, validation, deployment, operation, and monitoring.

- DOD should collaborate with private sector leaders on safety research in areas such as automated and autonomous driving systems, while prioritizing robustness and resilience research in areas understudied in the private sector.

- DOD should focus on traceability for assurance with ML systems that continue to learn on dynamic inputs in real-time.

- The United States should collaborate with allies on common standards for safety and security of AI systems, including AI-enabled safety-critical systems.

- The United States should pursue opportunities for collaboration with China and Russia on AI safety and maintain crisis communications protocols to reduce the risk of escalation.

Today's research and development investments will set the course for the future of AI in national security. For a more detailed analysis of the scope, nature, and strategic implications of U.S. military investments in autonomy and AI, we encourage our readers to turn to the two CSET reports: "U.S. Military Investments in Autonomy and AI: A Budgetary Assessment" and "U.S. Military Investments in Autonomy and AI: A Strategic Assessment."

Document Identifier: doi: 10.51593/20200070

Endnotes

[1] Congressional Research Service, *Federal Research and Development (R&D) Funding: FY2020* (Washington, DC: CRS, updated March 2020), 3, https://fas.org/sgp/crs/misc/R45715.pdf.
[2] Some of the recommendations in this report were also articulated in a report published jointly by the Bipartisan Policy Center and CSET. See Bipartisan Policy Center and the Center for Security and Emerging Technology, *Artificial Intelligence and National Security*, (Washington, DC: BPC, CSET, June 2020), https://bipartisanpolicy.org/wp-content/uploads/2020/07/BPC-Artificial-Intelligence-and-National-Security_Brief-Final-1.pdf.
[3] Congressional Research Service, *Defining Readiness: Background and Issues for Congress* (Washington, DC: CRS, June 2017), https://fas.org/sgp/crs/natsec/R44867.pdf.