Sailor aboard guided-missile destroyer USS *Michael Murphy* takes part in Office of Naval Research demonstration of new and improved training combining software and gaming technology to help naval forces develop strategies for diverse missions and operations, Pearl Harbor, March 24, 2016 (U.S. Navy/John F. Williams)

Intelligence in a Data-Driven Age

By Cortney Weinbaum and John N.T. Shanahan

In a foreseeable future, battles may unfold using weapons and tactics that the United States is ill-prepared to detect or counter. Today's ballistic missiles take tens of minutes to cross an ocean, but tomorrow's hypersonic weapons may take merely minutes. Urban warfare could occur in hyper-connected cities where overhead sensors provide limited value, while ubiquitous ground sensors provide too much data for analysts to mine. In the cyber domain, by the time an operator detects a "launch," a weapons package may have already reached its target and achieved its desired effect. Attacks against satellites, economic attacks, and covert influence campaigns can all occur undetectable to the human senses until too late.

The vector, volume, velocity, variety, and ubiquity of data are disrupting traditional tools and methods of national security policy, operations, and intelligence. The scope of such disruption will only grow and accelerate. Under the adage that "information is power," society has created technologies capable of creating volumes of structured and unstructured data so large as to overwhelm all previous forms of analytic tradecraft and pattern recognition. As part of their recommendations from the January 2017 public meeting, the U.S. Defense Innovation Board asserted that whoever amasses *and organizes* the most data—about ourselves as well as our adversaries—will sustain technological superiority.[1] Failure to treat data as a strategic asset will cede precious time and space to competitors or adversaries.

The U.S. Intelligence Community (IC), to include the Defense Intelligence Enterprise, faces daunting challenges

Cortney Weinbaum is a Management Scientist at the RAND Corporation. Lieutenant General John N.T. "Jack" Shanahan, USAF, is Director for Defense Intelligence (Warfighter Support) in the Office of the Under Secretary of Defense for Intelligence.

of pure scale—volume and velocity—as well as an ever-increasing complexity of data—variety and veracity. The IC is challenged to acquire, manage, correlate, fuse, and analyze ever-increasing amounts of data across agencies and with allies and partners. In our experiences, data in the IC are generated in too many diverse formats, in too many disconnected or inaccessible systems, without standardized structures and without overarching agreed-upon ontology. This situation risks wasted collections, lack of timeliness, missed indications and warnings, and lack of relevance for decisionmaking. The result is an inability to fuse data to create multi-sourced intelligence as early in the intelligence cycle and as close to the point of collection as possible. Analysts are given a task too difficult, too cumbersome, and with too many hurdles to clear to provide timely and relevant analytic judgments or actionable intelligence to policymakers and warfighters.

These challenges should be addressed by:

- embracing machine-learning algorithms that can parse data, learn from the data, and then respond
- encouraging creativity and deep thinking by intelligence professionals
- designing the policy, information technology (IT), agile acquisition, and security environment that allows human-machine tradecraft to flourish.

These problems cannot be solved within any one agency, program, or intelligence discipline. We see a compelling need for creative ways to adapt to this new environment that must include improving the technological and operational advantages of the IC with systems and machines capable of manipulating and understanding big data, as well as advancing human-machine and machine-machine collaboration, so analysts can make the best use of their time working on the hardest problems.

Meanwhile, serious questions linger about the unforeseen repercussions of a machine-learning "black box" that can generate solutions in ways that might not be readily explainable to its human operators. Artificial intelligence (AI) systems have created their own languages that human programmers could not read[2] and have taught themselves to play games using tactics that humans did not teach and cannot comprehend.[3] The repercussions of these effects in national security systems are unknown, untested, and remain largely unexplored.

The Future Battlespace

The future battlespace is constructed of not only ships, tanks, missiles, and satellites, but also algorithms, networks, and sensor grids. Like no other time in history, future wars will be fought on civilian and military infrastructures of satellite systems, electric power grids, communications networks, and transportation systems, and within human networks. Both of these battlefields—electronic and human—are susceptible to manipulation by adversary algorithms.

In electronic environments, algorithms are already used to monitor and maintain control over most areas of critical infrastructure (electric, water, food, financial, communications, and so forth). Russia and China have demonstrated their interest in testing the capabilities and weaknesses of these systems in the United States, and intelligence agencies need the ability to fuse data across multiple sources to understand adversary activities and intended outcomes.

To disrupt human networks, the theft of personal data on cleared government workers in the Office of Personnel Management breach provides a rich data set for an adversary to tailor a covert influence campaign against *each individual* military leader or policymaker.[4] If this data were to be combined with financial records stolen from Equifax, email records from Yahoo!, medical information from Anthem health insurance, and data from additional sources, algorithms could create highly sophisticated and individualized covert influence campaigns against the United States. In a less sophisticated campaign, North Atlantic Treaty Organization military forces recently reported that soldiers' phones were hacked by Russia during military training exercises "to gain operational information, gauge troop strength, and intimidate soldiers," according to Alliance officials.[5]

The ability to fuse enormous amounts of data from across disparate data sets and provide meaningful answers are exactly what artificial intelligence and machine learning were designed to do. As long as the commercial sector can find a way to use data to anticipate the brand of car or toothbrush that a consumer is likely to purchase, vendors will sell capabilities to identify user preferences and weaknesses with a specificity that intelligence officers might expect to find in psychological profiles.

The United States is at risk of allowing adversaries to accelerate and steal the competitive advantage. China has a national strategy for AI with commensurate pledges to invest billions of dollars in AI technologies over the next 5 years.[6] Chinese researchers publish more journal articles on AI than their U.S. counterparts,[7] and People's Liberation Army strategists are preparing for a world where humans cannot keep pace with battlefield decisionmaking.[8]

In the United States, the recently published National Security Strategy and National Defense Strategy both address the importance of AI and autonomy to national security and warfighting.[9] Beyond the 2016 National Artificial Intelligence Research and Development Strategic Plan, however—which is largely focused on research and development—the United States does not yet have a sweeping national strategy for AI.

Eric Schmidt, executive chairman of Alphabet (parent company of Google) and chair of the Defense Innovation Advisory Board, described China's advances in AI compared to the United States: "By 2020, they will have caught up. By 2025, they will be better than us. By 2030, they will dominate the industries."[10]

A Looming Intelligence Failure

Future intelligence tradecraft will depend on accessing data, molding the right enterprise architecture around data, developing AI-based capabilities to dramatically accelerate contextual understanding of data through human-machine and machine-machine teaming, and growing analytic expertise capable of swimming and navigating in enormous data lakes. The IC needs to develop tradecraft and methodologies for accessing, arranging, and analyzing data, including structured analytic techniques and analytic tradecraft standards for machine intelligence. New technology is evolving faster than the ability of the Department of Defense (DOD) and IC to implement it, train on it, and use it effectively.

Within the Defense Intelligence Enterprise, investments in collectors and sensors are generating an ability to collect more data from more sensor types than at any time before. The DOD roadmap for unmanned systems describes a plan for thousands of unmanned air, sea, and ground systems, without a clear path for how all the data from those systems will be analyzed to create value.[11] This is in addition to space systems and publicly available unclassified systems. An increase in collection, to include from the most highly classified exquisite sensors, does not necessarily equate to more or better intelligence or orientation, especially when facing near-peer competitors who may prove equally adroit at adapting to the information environment.

DOD Project Maven, led by Lieutenant General Shanahan, has a goal of overcoming human intelligence analysts' inabilities to deal effectively with the massive amounts and types of collection across every domain, a quandary called "success catastrophes."[12] As a starting point and pathfinder project, former Deputy Secretary of Defense Robert Work tasked the Maven team to find AI and computer vision solutions to augment, amplify, and automate exploitation of unmanned aerial system full-motion video.

Current intelligence practices involve extracting information of value from large datasets of cross discipline (cross-intelligence) information—the needle in the haystack—leading to the bulk collection and storage of hay in hopes that eventually all needles will exist inside. In a more data-oriented era, it is increasingly possible to draw intelligence of value from the data in aggregate (temporal and geospatial behavior patterns, for example). This can result in an ironic dilemma in which there is too much data for humans to search effectively for needles, yet not enough accessible data from which to draw and validate useful intelligence.

Next is the question of what to do with intelligence once it is attained. The military Services are developing and acquiring combat systems with a greater hunger for data and *intelligence* than in the past, and intelligence mission data must be transferred to these systems as early as possible in the acquisition cycle and then updated frequently and fast enough for use in combat, in ingestible data structures, and at classification levels the combat systems can handle.[13] Meanwhile, within the policy community, policymakers are increasingly relying on unclassified publicly available information when classified intelligence is too slow to arrive or too highly classified to be useful.[14]

In his groundbreaking work on the *observe*, *orient*, *decide*, and *act* (OODA) loop, Colonel John Boyd, USAF (Ret.), emphasized the importance of operating at a tempo or rhythm that an adversary cannot comprehend or match. Operating inside an adversary's OODA loop helps accomplish those objectives by disorienting or warping an opponent's mental images so that he can neither appreciate nor cope with what is happening around him. In today's fast-paced and ever-changing data-driven age, the terms *information dominance* or *information superiority* are chimerical; instead, *temporal advantage* might be the best possible outcome. Yet even that could be sufficient to gain the upper hand, if U.S. warfighters can stay inside the adversary's OODA loop while simultaneously using data in imaginative ways to distort the adversary's own orientation.

Artificial intelligence and machine learning provide opportunities to accelerate through every step of the OODA loop by making sense of data in real time as the data arrive, evaluating options and initiating an action in milliseconds, and acting. Such decisions may include responding to indications and warnings before human operators have time to read an alert or initiating a response within a predetermined set of approved parameters. Machine learning offers new opportunities to shrink the first two phases of the OODA loop, greatly increasing the potential for humans to accelerate decisionmaking and taking action.

We may well be facing a future involving algorithm-versus-algorithm warfare, leading us to question whether 21st-century warfighters might look at minutes of decision time as luxurious relics of the past.

Solutions Exist Within Reach

Intelligence agencies can and should invest in cross-domain, cross-program, and cross-discipline machine-learning capabilities and require intelligence officers to use these capabilities to their fullest potential. Any data that remain stovepiped in compartments, proprietary databases, and on classified domains that algorithms cannot reach will require manual integration by intelligence officers, delay intelligence assessments, and create protected bubbles of data where officers may not be able to see inside to bring all sources to bear on analytic problems. This vision threatens concepts of "need to know" and would force the collection and analytic communities—with their brethren in counterintelligence and security offices—to reconcile threats from outside with threats from within. Data protection policies may give algorithms access to data fields that human analysts are not cleared to see, possibly requiring decisions about how much trust can be placed in machines and how their work can be audited for vulnerabilities that are both naturally occurring and adversary generated.

To create this endstate, several activities should be considered. First, finding the answer to any intelligence question should start with the proposition that every analyst needs all potentially relevant data, from every possible source.[15] This suggests striking a different balance between the classic deductive (searching for the known unknowns) and inductive

Soldier with 3rd Battalion, 6th Field Artillery Regiment, 1st Brigade Combat Team, 10th Mountain Division (LI), navigates new Precision Fires-Dismounted system, which includes viewing live-streaming full-motion video from unmanned aerial vehicles on smartphone, 3D digital maps, and ability to send precision target coordinates, at Mission Training Center, Fort Drum, April 5, 2018 (U.S. Army/James Avery)

(synthesizing to discover the unknown unknowns) analytic approaches. It also requires a different approach to collection because all data may be relevant long after collection and should be accessible in discoverable archives; the processes for doing this will depend heavily on whether datasets include information on U.S. persons and other protected entities, while data controls and data quality assurance become essential functions.

Widespread integration of machine learning and AI will present new opportunities for deception resulting from data that have been altered or manipulated. Counter-AI will become prevalent while influence operations will take on new dimensions that have yet to be fathomed, requiring a renewed emphasis on both offensive and defensive cognitive-centric operations. Intelligence analysts need to be trained on how to recognize attempts by an adversary to use altered or manipulated data, including understanding how to use AI to maximum advantage to prevent even the more sophisticated influence operations from affecting desired operational outcomes.

Second, data would not be treated as an IT problem; instead, IT systems should be framed by the operational problems they solve. This requires moving from closed, proprietary architectures and untenable lack of data standards to open architecture and Agile Methodology—open architectures and fast transient adoption of new technologies and applications—where any data from any source can be found and ingested by any analyst at any time. More often, algorithms will be moved to the data, rather than trying to move data to the algorithms. Global cloud solutions are essential to integration, optimized for all aspects of AI rather than only for data storage or search. Data access must be mastered to provide the fuel for machine learning and human-machine teaming. In turn, rapid data access requires effective data management, which calls for new skill sets and expertise—such as data architects and data scientists. Network access across all security domains, access to all relevant data types, and agile integration of disruptive technologies are key to achieving and sustaining decision advantage.

Third, publicly available information and open source information will provide the first layer of the foundation of our intelligence knowledge. This requires a major shift from assuming that the highest classified intelligence is the most infallible to embracing and integrating nontraditional and unclassified sources. Exquisite collection from all other intelligence disciplines will enhance foundational intelligence and fill in existing knowledge gaps. This flips a 60-year paradigm and challenges the very concept of "intelligence"

Airman with 379th Operations Support Squadron performs maintenance on satellite dish at Al Udeid Air Base, Qatar, March 30, 2018, as part of Operation *Silent Sentry*, protecting critical satellite communication links (U.S. Air National Guard/Phil Speck)

when classification is not a requirement for information to be of intelligence value.[16]

Fourth, shift the joint and combined analytic workforce from industrial-age production line processing and exploiting single collection streams of data to an information-age enterprise model where some analysts conduct multi- and all-source correlation and fusion, fully integrated with joint, national, and international partners. While layering intelligence data is a decent start, it is insufficient. Both human and AI system sense-making are needed to deliver time and space for decisionmaking. This principle also introduces broader questions about the future balance of breadth and depth across the analytic workforce. Training would require more emphasis on synthesis and creativity in analysis.

Finally, the solutions described above would require a revolution in IC life cycles for human capital, budgeting, acquisition, and research and development. Hiring and security clearance processes that last 2 years or more result in agencies on-boarding employees who were at the top of their game 2 years ago (an eon in a data-driven world), and, for mid-career hires, position requirements value experience in government over science, technology, and analytic experience in the commercial and academic sectors.[17] In the best of circumstances DOD and the IC have been challenged to create multiyear budget strategies within the 4-year Planning, Programming, Budgeting, and Execution process, while in today's climate of chronic continuing resolutions, creating a budget *strategy* is a hope rather than a regular occurrence. Intelligence agencies have tried again and again to create innovative acquisition reforms using small subsets of their budgets and "innovation offices," but these solutions stall when scaled across national or military intelligence programs.

Changes to the IC's traditional acquisition processes will require a generation of contracting officers who have the training and resources to manage an overhaul of contracting processes that puts focus back on quality, results, and the speed of relevancy rather than defaulting to lowest price technically acceptable contracts.[18] Adapting Agile Methodologies would facilitate faster paces of technology development, implementation, and refinement. Finally, each of these reforms would create an environment where research and development (R&D) offices—in IC agencies and military Services—can thrive. R&D organizations need the brightest technologists to build partnerships among collectors, analysts, and industry vendors, and they need the support of proactive contracting officers and an effective budget environment to succeed—ultimately leading to an AI-ready, *prototype warfare* culture.

Concluding Thoughts

Our proposal has at least one Achilles' heel that the United States should plan for and mitigate: an over-reliance on technology. Even in the age of autonomous systems, war will remain a human endeavor. If the Nation were to fight a technologically primitive enemy, such as in the mountains of Afghanistan or jungles in Africa, warfighters and intelligence officers risk being too reliant on systems that require large quantities of data. Alternatively, in a near-peer scenario the United States may one day fight an enemy who finds and exploits vulnerabilities in our technology and blinds our warfighters or uses data against us in new and creative ways. As a result of both scenarios, the Nation will continue to value intelligence analysts and warfighters skilled in low-tech tested and reliable tradecraft and solutions.

The best intelligence analysis derives from the right combination of art and science. The art of intelligence may be the same today as it was 2,000 years ago. What is different now, however, is the necessity of getting much better much faster at the science of the tradecraft, which is centered on data. Analysts must have the tools they need to deal with massive amounts of information that enable them to close intelligence gaps and enable better operational outcomes at the speed of data.

Artificial intelligence and machine learning will be instrumental to increasing the effectiveness of the future intelligence analyst workforce, improving the odds of gaining and sustaining a competitive or temporal advantage. Digital transformation, methodic multidomain data integration, and algorithmic warfare will be the heart of the intelligence enterprise's role in sustaining a long-term competitive advantage. This is as much about strategic innovation as it is innovation at the tactical or analyst level. One without the other is necessary but insufficient.

The IC is nearing critical decisions on AI and machine learning. Despite a number of disadvantages inherent in 16 years of continuous counterterrorism and counterinsurgency operations, the IC forged a highly experienced, battle-trained analytic workforce unlike any other in the world. The IC's greatest potential asymmetric strengths remains its ability to make sense of data quickly and remaining inside the adversary's OODA loop. Will the United States and its adversaries slow down their machines to the speed of human thought to maintain a man in the loop? Or will each country pursue AI to its fullest potential, fearing that if not, its adversaries will pursue it first? The time has arrived for the Intelligence Community to decide how it wants to answer these questions. JFQ

Notes

[1] "Defense Innovation Board Holds Public Meeting," *DOD News*, video, 1:54:30, October 24, 2017, available at <www.defense.gov/Videos/videoid/560180/#DVIDSVideoPlayer1133>. Emphasis added.

[2] Mark Wilson, "AI Is Inventing Languages Humans Can't Understand. Should We Stop It?" *Co.Design*, July 14, 2017, available at <www.fastcodesign.com/90132632/ai-is-inventing-its-own-perfect-languages-should-we-let-it>.

[3] Larry Greenemeier, "AI versus AI: Self-Taught AlphaGo Zero Vanquishes Its Predecessor," *Scientific American*, October 18, 2017, available at <www.scientificamerican.com/article/ai-versus-ai-self-taught-alphago-zero-vanquishes-its-predecessor/>.

[4] Sina Beaghley, Joshua Mendelsohn, and David Stebbins, "What Is the Adversary Likely to Do with the Clearance Records for 20 Million Americans?" *Inside Sources*, January 19, 2017, available at <www.insidesources.com/adversary-likely-clearance-records-20-million-americans/>.

[5] Thomas Grove, Julian E. Barnes, and Drew Hinshaw, "Russia Targets NATO Soldier Smartphones, Western Officials Say," *Wall Street Journal*, October 4, 2017, available at <www.wsj.com/articles/russia-targets-soldier-smartphones-western-officials-say-1507109402>.

[6] Graham Webster et al., "China's Plan to 'Lead' in AI: Purpose, Prospects, and Problems," *New America*, August 1, 2017, available at <www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/>.

[7] "China May Match or Beat America in AI," *The Economist*, July 15, 2017, available at <www.economist.com/news/business/21725018-its-deep-pool-data-may-let-it-lead-artificial-intelligence-china-may-match-or-beat-america>.

[8] Elsa B. Kania, "Artificial Intelligence and Chinese Power," *Foreign Affairs*, December 5, 2017, available at <www.foreignaffairs.com/articles/china/2017-12-05/artificial-intelligence-and-chinese-power>.

[9] *National Security Strategy of the United States of America* (Washington, DC: The White House, December 2017), available at <www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf>; *Summary of the National Defense Strategy: Sharpening the American Military's Competitive Edge* (Washington, DC: Department of Defense, 2018), available at <www.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>.

[10] Patrick Tucker, "China Will Surpass U.S. in AI Around 2025, Says Google's Eric Schmidt," *Defense One*, November 1, 2017, available at <www.defenseone.com/technology/2017/11/google-chief-china-will-surpass-us-ai-around-2025/142214/>.

[11] *Unmanned Systems Integrated Roadmap FY2013–2038* (Washington, DC: Department of Defense, 2013), available at <http://archive.defense.gov/pubs/DOD-USRM-2013.pdf>.

[12] Cheryl Pellerin, "Project Maven to Deploy Computer Algorithms to War Zone by Year's End," *DOD News*, July 21, 2017, available at <www.defense.gov/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/>.

[13] *Unmanned Systems Integrated Roadmap FY2013–2038.*

[14] Derek Grossman, "Keeping Up with the Policymakers: The Unclassified Tearline," *War on the Rocks*, July 28, 2016, available at <https://warontherocks.com/2016/07/keeping-up-with-the-policymakers-the-unclassified-tearline/>.

[15] *The 9/11 Commission Report* (Washington, DC: National Commission on Terrorist Attacks Upon the United States, 2004), available at <www.9-11commission.gov/>.

[16] Rich Girven, Sina Beaghley, and Cortney Weinbaum, "A New Paradigm for Secrecy," *U.S. News & World Report*, October 13, 2015, available at <www.usnews.com/opinion/blogs/world-report/2015/10/13/defining-a-new-paradigm-for-government-secrecy>.

[17] Lindy Kyzer, "Top Secret Clearance Processing Times Lengthen, Secret Times Improve," *ClearanceJobs.com*, November 1, 2017, available at <https://news.clearancejobs.com/2017/11/01/top-secret-clearance-processing-times-lengthen-secret-times-improve/>.

[18] Scott R. Calisti, "Lowest Price Technically Acceptable: Why All the Debate?" *Defense AT&L*, March–April 2015, available at <http://dau.dodlive.mil/2015/03/01/lowest-price-technically-acceptable-why-all-the-debate/>.