OPEN-SOURCE DATA IS EVERYWHERE—EXCEPT THE ARMY'S CONCEPT OF INFORMATION ADVANTAGE

Maggie Smith and Nick Starck | 05.24.22

Editor's note: This article is part of the series "Compete and Win: Envisioning a Competitive Strategy for the Twenty-First Century." The series endeavors to present expert commentary on diverse issues surrounding US competitive strategy and irregular warfare with peer and near-peer competitors in the physical, cyber, and information spaces. The series is part of the Competition in Cyberspace Project (C2P), a joint initiative by the Army Cyber Institute and the Modern War Institute. Read all articles in the series here.

Special thanks to series editors Capt. Maggie Smith, PhD, C2P director, and Dr. Barnett S. Koven.

Three months ago, as Russia invaded Ukraine, the world watched as Twitter exploded with real-time data, reporting, and analysis of the unfolding conflict. It quickly became clear that the war presented analysts with an unprecedented amount of rich, open-source data on military movements, troop location, shelling damage, weapon types, and more. Ukraine has been quick to capitalize on Russia's poor data protection and President Volodymyr Zelenskyy has become Ukraine's most potent weapon because of his ability to use data and information and Russia's inability to protect it.

For the US Army, a key takeaway from the Ukrainian conflict so far should be the extent to which our modern-day habits are trackable, traceable, and predictable. Open-source data presents modern militaries, especially wealthy high-tech ones, with a very uncomfortable truth: militaries are exposed because their troops are connected. Currently, the US legal and regulatory systems do not, and cannot, protect the average citizen—and therefore, the average US service member—from risks associated with the ubiquitous open-source data produced by our surveillance economy. From a national security perspective, the accumulation of open-source data on people—their habits, their likes and dislikes, their exercise routines, and more—and its potential to impact the military's ability to fulfill its man, train, and equip mandate from Congress is deeply concerning. Also alarming is the amount of information our adversaries can glean about US strategic interests from tracking US military activity on any number of apps, like Flightradar24, which includes US military reconnaissance platforms such as the unmanned RQ-4 Global Hawk, the RC-135V Rivet Joint, and others among the aircraft it tracks, and Strava, the fitness tracking app. Ultimately, you can intuit quite a bit about where our forces may be heading, where military planners are focusing their efforts, and where the next conflict is likely to occur if you simply track where Rivet Joints are conducting sorties and service members

Department of Defense.

MOST POPULAR POSTS

Underground Nightmare: Hamas Tunnels and the Wicked Problem Facing the IDF

What Can the IDF Do about Hamas Tunnels?

are working out. And for the Army specifically, the existing and emerging doctrine fails to account for the surveillance economy and its open-source data, leaving a gaping hole in our competitive strategy.

Information Advantage: What Is It?

Presently, the Army is developing its doctrine for its newest term of operational art: information advantage. Information drives friendly, neutral, and adversary actors at all levels and across all domains of warfare. Information advantage is a condition of relative advantage that enables a more complete operational picture and leads to decision dominance—the sensing, understanding, deciding, and acting faster and more effectively than the adversary. Gaining the initiative and maintaining a position of relative advantage over the information environment—regardless of where we find ourselves on the conflict continuum—largely depends on a commander's ability to achieve an information advantage over a defined target audience or adversarial decision maker in a specific context or timeframe. Complementary to information advantage is the employment of information and other capabilities as weapons, designed to shape friendly, neutral, and adversarial perceptions, attitudes, and behaviors. Ultimately, the ability to shape perception and achieve victory in modern conflict and competition is heavily dependent on trust—trust in data, among team and unit members, in leaders, in doctrine, in equipment, and in capabilities.

To achieve information advantage, the Army conceives of five, interrelated core tasks— what have been described as "information advantage activities." Commanders must: 1) enable decision making; 2) protect friendly information; 3) inform and educate domestic audiences (a task conducted in accordance with laws and focused on public affairs office activities); 4) inform and influence international audiences; and 5) conduct information warfare. In theory, information advantage activities are synchronized through the operations process, integrated

across the Army's six warfighting functions—command and control, intelligence, protection, movement and maneuver, fires, and sustainment—and employed using all available military capabilities. After distilling the Army's rhetoric, information advantage requires commanders to prioritize persistent sensing, ongoing analysis, cyclical assessments, and a willingness to continuously update assumptions to ensure they maintain a dynamic situational awareness of the environment—in competition and conflict. Ultimately, the Army anticipates that victory in future warfare, and in the current era of persistent engagement, will come down to who can gain the most by effectively employing information to their advantage.

The National Security Risk

Instead of gunfire or artillery explosions, some of the first signs that Russia was invading Ukraine on February 24, 2022, came from Twitter. For example, Dr. Jeffrey Lewis, an expert in arms control and nonproliferation, compiled open-source data from the traffic layer of Google Maps and shared the Russian troop movements he identified, essentially in real time, on Twitter. According to Google Maps, he Tweeted, "there is a 'traffic jam' at 3:15 in the morning on the road from Belgorod, Russia to the Ukrainian border"—exactly the spot where vehicles, equipment. and manpower had massed the previous day. "Someone's on the move," Dr. Lewis concluded, and he was right. As the Ukrainian conflict escalated, individual researchers and organizations continued to collect and analyze open-source data—also defined as publicly available information by DoD—from social media platforms, commercial satellites, and public databases. Their analysis and reporting have emerged as a critical resource on the conflict, providing combatants and observers with incredible insight and minute-by-minute assessments of what is happening on the ground.

However, the ability to track ongoing military operations through open-source data is not new—in 2016 Bellingcat released a report that used open source data to

document the full scale of the Russian artillery attacks against Ukraine in the summer of 2014. In fact, using open-source data is the new normal. And various US government agencies, including the Department of Defense, rely on open-source data for intelligence and procure data through contracts with data brokers. In response, civil society and privacy watchdogs around the world have voiced concern, highlighting the risks to personal privacy associated with government-led data collection, aggregation, and use. The likely result is new legislation, like the proposed Fourth Amendment is Not For Sale Act and others.

However, the use of open-source data and large scale, legal data collection efforts frequently pose less obvious national security risks. China, for example, aggressively collects data—legally and illegally—to support its domestic and international goals. A major threat to US citizen data is China's Beijing Genomics Institute (BGI), which has grown into one the world's largest genomic companies after working on the Human Genome Project. BGI developed a prenatal genetic test, in collaboration with the Chinese military, that is sold and used globally. However, in addition to providing prospective parents with important genetic information, the DNA specimens are also amassed into a vast bank of genomic data that China is using to conduct large-scale studies of population traits. More than eight million women have taken BGI's prenatal tests globally, and China has their DNA and location data stored locally in mainland China. BGI also developed a COVID-19 test and offered to set up testing laboratories in several US states at the start of the pandemic. Mike Orlando, head of the National Counterintelligence and Security Center, identified the BGI offers as a national security risk, "citing concerns about how China might use personal data collected on Americans." Even when done legally, DNA collection by Chinese companies should be understood as part of China's comprehensive effort to collect records and data.

On the other hand, data also creates risk for the governments that aggressively pursue it. Experts are

increasingly identifying the ways that open-source data can be used to expose government activity (e.g., military maneuvers, resource allocation, travel, or policy activity) and how the ever-growing pools of open-source data generated by modern societies pose a national security risk. But we lack precision in how we describe the sources, mechanisms, and outcomes of open-source data risks, preventing the development of a coherent mitigation strategy tailored to the national security context. Without a common understanding of risk, civilian and military leaders are unable to make informed and consistent decisions about open-source data, leading to strategic missteps and tactical knee-jerk reactions—like embedding code in the Free Application for Student Aid website that sends user information back to Facebook (the code has since been removed) or banning service members from using geolocation features on devices in deployed areas (e.g., fitness trackers).

What Is Information Advantage Missing?

The piece missing from the Army's information advantage framework is an awareness of how the persistent aggregation of open-source data in the surveillance economy impacts the Army's ability to achieve information advantage. Because the American public is subject to the surveillance economy, US service members are, too. George Washington famously emphasized that "when we assumed the Soldier, we did not lay aside the Citizen" as a cautious reminder that soldiers are citizens first. Since service members live alongside and among the general population, service members and veterans are not only susceptible to the same targeted marketing the average citizen experiences, but are actually the target of additional foreign manipulation and surveillance efforts. Soldiers, sailors, airmen, and Marines access social media platforms and online services just as civilians do. They also purchase items online, apply for credit cards online, do their taxes with online tax preparation tools, and surf the web just like their civilian counterparts. But unlike the civilian neighbors they barbeque with, they also fight the nations' wars. Open-source data is produced continuously

by all service members as they go about their digitally connected lives alongside their civilian counterparts, making the surveillance economy an integral part of the information environment that commanders need to consider as they conduct information advantage activities.

The military is beginning to understand the potential risks presented by open-source data, particularly in combat situations, partly because examples of how open-source data can expose military information abound—from troop location tracking on Tinder to tracking stolen AirPods to SIM cards revealing Russian troop locations in Ukraine. Of course, these known cases fit neatly within traditional operational security risks and are scenarios that senior military leaders can relate to—especially when open-source data is directly contributing to deaths on the battlefield or to the identification of war criminals. However, having a tactical appreciation of the open-source data risks during periods of declared conflict is not enough to achieve information advantage—the risks to military operations are present well before any decision to go to war is made, and persist after conventional conflict ends. In fact, the risks are a constant factor in the current competition environment, making any ex post facto restrictions, regulations, or rules placed on deployment behavior inadequate and misguided. Changes need to happen at home, well before the deployment cycle begins. Failing to consider garrison operations and the ways that soldiers interact with the surveillance economy as part of the information environment that commanders need to consider for information advantage is a failure to understand when and where the vulnerabilities and threats to the force begin and a failure to account for our modern, digitally connected, human behavior.

The "So What" of Open-Source Data

For multi-domain operations, the Army frames the operating environment as including human, physical, and informational aspects. To be effective across the

competition continuum, the Army proposes positioning formations and capabilities forward, so that information advantage activities are integrated into security cooperation efforts and crisis action planning on behalf of theater commanders. To coordinate information advantage activities in an area of conflict, the Army identifies that preparation must begin in competition, or when forces develop the intelligence to identify specific vulnerabilities and then gain or prepare to request the required authorities, and train to use national-level capabilities. The overall goal is operational convergence with formations postured to degrade, disrupt, or destroy adversary capabilities, while defending those of friendly forces. However, what this framework does not consider is the intersection of the human, physical, and informational aspects, or the risks to day-to-day garrison operations from open-source data.

Ultimately, the risks of open-source data are not an individual's problem, but an Army problem. For example, fake accounts on Facebook for US Army general officers are numerous, and in some cases, fail to violate Facebook's terms of service and can therefore, remain active. Even LinkedIn is rife with fake profiles attempting to make connections with users in targeted marketing campaigns. Additionally, fake social media accounts managed by Russia have already mobilized the American public in connection with divisive issues, making fake accounts for authoritative figures, like US Army generals, especially concerning. From a national security perspective, open-source data enables foreign manipulation efforts that target the US military and veteran populations through the use of "misleading and divisive questions about the U.S. government's military and veteran policies to further amplify and exploit the existing frustrations." The relative ease with which anyone can purchase open-source data means that soldier data is already being used to target service members for products, media, or other services and presently, there is nothing preventing our adversaries from using open-source data to target them as well.

To achieve information advantage, the Army needs to give commanders the tools necessary to assess the operational risks of open-source data, social media, and related information technologies. The Army has longstanding doctrine for assessing operational risks; however, the traditional risk management framework is intentionally broad, leaving commanders without clear guidance or terminology for identifying, assessing, and making risk decisions in the information environment. As the Army develops its information advantage doctrine, it should simultaneously develop a dedicated data risk management framework to enable modern commanders to achieve information advantage. In its current form, information advantage perpetuates an antiquated notion that operating environments are (or can be) geographically bound—as the conflict in Ukraine has highlighted, kinetic actions may be limited to a geographic area, but informational risks are global. A dedicated data risk management framework would be a guide for commanders to continually and methodically assess the evolving information environment, to identify and address conceptual gaps, and to achieve their informational and operational goals. As the information environment emerges as the main effort in competition and conflict, the Army must adapt and provide its commanders with the right concepts, doctrine, and resources to succeed in a world characterized by the ubiquity of open-source data.

Captain Maggie Smith, PhD, is a US Army cyber officer currently assigned to the Army Cyber Institute at the United States Military Academy where she is a scientific researcher, an assistant professor in the Department of Social Sciences, and an affiliated faculty of the Modern War Institute. She is also the coeditor of this series and director of the Competition in Cyberspace Project.

Captain Nick Starck is a US Army cyber officer currently assigned as a research scientist at the Army Cyber Institute.

His research focuses on information warfare and data privacy.

The views expressed are those of the authors and do not reflect the official position of the United States Military Academy, Department of the Army, or Department of Defense.

Image credit: Sgt. Dustin D. Biven, US Army