North Carolina Army National Guardsman with Detachment 4, Recruiting and Retention Battalion, Joint Force Headquarters, exits simulated shoot house, pulling litter with simulated casualty, as competitor in 2013 Army National Guard Best Warrior Competition, Little Rock Arkansas, July 23, 2013 (U.S. Army/Betty Boyce)

Continuing the Big Data Ethics Debate

Enabling Senior Leader Decisionmaking

By Paul B. Lester, Pedro S. Wolf, Christopher J. Nannini, Daniel C. Jensen, and Delores Johnson Davis

Major Paul B. Lester, Ph.D., USA, is Director of the Research Facilitation Laboratory, Office of the Deputy Under Secretary of the Army (ODUSA). Pedro S. Wolf is a Behavioral Scientist in the Research Facilitation Laboratory, ODUSA. Christopher J. Nannini is Program Manager at the Research Facilitation Laboratory, ODUSA. Daniel C. Jensen is Director of the Army Analytics Group, ODUSA. Delores Johnson Davis is the Senior Professional for Integration (Human Dimension) in the Office of the Assistant Secretary for Manpower and Reserve Affairs.

In *Joint Force Quarterly* 77, Karl Schneider, David Lyle, and Francis Murphy presented a foundational debate on the ethical use of big data within a military context. The authors offered several cases where the military would benefit from improving its analytical capabilities to leverage the potential that big data offers. Most germane to the current article, they argued that "the collection and use of big data cannot compromise the organization's core value of trust: that the military will both provide for the national defense and also look out for the best interest of its Servicemembers."[1]

It would appear that their concern was quite prophetic, as recent data breaches in the private sector, government, and military continue to shed light on the endemic challenges that persist in the ethical use and protection of large volumes of sensitive personal data.[2] Though some have claimed that the era of privacy is waning,[3] both Federal law[4] and Department of Defense (DOD) policy[5] clearly state otherwise. Thus, we concur with Schneider, Lyle, and Murphy that it is incumbent on DOD leadership to safeguard personal data—the digital representation of a Servicemember's military history—within its system and use the data ethically both for research and policymaking. Yet a key question remains: How do we balance protecting the best interests of Servicemembers and maintaining their trust while also using available data and advanced analytics for the good of the Defense Department? This is a question that each Service must answer.

In the 2 years since *JFQ* published the big data ethics article, a multidisciplinary group of leaders within the Department of the Army has worked toward answering this question. The Federal Government and DOD have exercised tremendous leadership to balance privacy management with big data technology and training. DOD agencies now have an opportunity to consolidate data centers and systems to reduce the number of disparate silos and fuse data results for analytics-based projects and decisionmaking. Likewise, information technology systems with strong governance processes have emerged that place ethical, legal, and moral considerations at the forefront of approving personnel and medical data analytic projects. When coupled with recent big data policy decisions made within the Army Secretariat, current advances in this domain suggest that our military is at a critical policy juncture, presenting us with an opportunity to extend the debate on the ethical, legal, and moral use of Servicemember big data, as well as ensuring that DOD keeps pace with private-sector big data analytics innovations.

In this article, we begin by exploring the Human Capital Big Data (HCBD) initiative, an approved strategic policy framework intended to integrate and coordinate the ethical use of the Army's massive data stores by the research and analysis community. Next, we introduce the Person-Event Data Environment (PDE), the operational information technology platform designed from the ground up with the ethical use and security of Army personnel big data in mind. Later, we highlight some of the critical machine learning and predictive analysis research already under way within the PDE that supports the Army's personnel, medical, and intelligence communities. We close by outlining key operational and strategic opportunities and challenges of applying this emerging technology to the dynamic and complex human behavior we will face.

The HCBD Initiative

Like other Services, the Army has existing data stewardship strategy and policy that broadly sets data goals and governs the management, storage, and security of its data; this strategy is known as the Army Data Strategy (ADS), while the accompanying policy is known as the Army Data Management Program (ADMP). Though the ADS and ADMP present a comprehensive approach for strategic management of the Army's data and its information architecture, both documents take a neutral position toward unique characteristics of data and associated ethical, legal, and moral considerations. Thus, the Assistant Secretary of the Army (Manpower and Reserve Affairs) established a multidisciplinary working group in 2014 to begin addressing operational, security, and ethical considerations related to the use of big data in the human capital domain. A "big tent" approach was taken as members from major commands, information technology, military and civilian personnel, medical, training, legal, law enforcement, marketing, and research communities were invited to participate.

What first emerged from this working group was a white paper that accomplished three objectives. First, the paper described how new data policy related to the Human Capital Enterprise (HCE) would need to be nested within the Army's existing data stewardship policies, but that new HCE data policy terminology (taxonomy) should be harmonized with existing policies where possible. In short, the working group members concurred that emerging HCE big data policy would be a new branch growing on the larger Army Data Management tree. Second, the paper delineated how human capital data fit in a separate legal and ethical category from other data collected and stored by the Army, and further outlined how the proposed use of human capital data should trigger deliberate ethical, legal, and moral considerations. Quite simply, both the Privacy Act of 1974 and Health Insurance Portability and Accountability Act of 1996 set a high bar for organizational use of personal information, which in turn should be used as a guide during strategic planning for the legally acceptable use of human capital data. Third, the paper identified five fundamental principles DOD must consider in any future HCE big data endeavor, described in greater detail in the table.

The three objectives met via the white paper set the conditions for strategic planning intended to set key objectives for the use of big data in the human capital domain. Published in 2016, the Human Capital Big Data (HCBD) Strategy echoed many of the strategic goals found in the ADS and ADMP, particularly that data should meet VAUTI standards:

- *visible* by posting them to shared spaces and registering metadata related to structure and definition
- *accessible* to authorized users through those shared spaces and data services; will be controlled in accordance with the asset's security-related metadata
- *understandable* by creating data models, integrating data, and identifying requirements for information traceability
- *trusted* by identifying authoritative sources and making data storage and access structure
- *interoperable* by complying with information exchange specifications and establishing master data management and unique identifiers, which will allow the same data to be used across multiple systems and applications.

Table. Five Fundamental Principles of Army Big Data Policy

Transparency	Individuals are entitled to understandable information about how the Army collects data on them, who has access to that data, and how that data will be used and secured. A responsible enterprise approach must balance the tradeoffs made among privacy, security, and convenience.
Privacy	An individual's right to privacy is fundamental. A breach of privacy can become a breach of trust between the organization holding an individual's data and that individual, regardless if harm occurs. Collection of large amounts of data specific to an individual—even without the inclusion of personally identifiable information—cannot be assumed to maintain an individual's anonymity.
"Do no harm"	All necessary steps will be taken by the Army to ensure that application and use of data maximizes benefits and minimizes harm to Army personnel, individually and collectively.
Validity and verification	Consequential or preemptive prediction applications of data will be held to accepted scientific standards of validity and verification with appropriate peer review before implementation within the Army.
Security	Datasets must be protected from both internal and external threats. This maintains the fidelity of the data and keeps faith with our people. Users access to Big Data, particularly as datasets are combined and stored together, needs to be specifically addressed.

The HCBD Strategy also tackles several of the operational challenges discussed by Schneider, Lyle, and Murphy by outlining six guiding principles that should be practiced within the HCE. First, the HCBD data management culture must exist in a manner that reinforces trust within the Army culture. While the Service has a duty to share data when legally permissible, this legal requirement should be balanced with the notion that Army personnel must also have confidence in the accuracy, secure storage, and ethical and legal use of the data. Second, the quality of any HCBD endeavor largely depends on systems engineering, human-system integration, and user training. In most cases, high-quality data stem from having information technology platforms that are engineered with traceable data quality metrics and objectives and are relatively easy to use by trained personnel. Third, a common lexicon and taxonomy are necessary to create an operating vocabulary for shared situational understanding and transparency across the diverse silos of data. For example, a common term used in the Army's data environment is *data owner*, which suggests that an individual or organization managing data may make final decisions about when and how the data may be shared. Yet in most cases, individuals or organizations are actually *data stewards* charged with the collection, management, and operational use of the data, leaving decisions on data-sharing to be made by a higher authority. Thus, developing a common lexicon helps to standardize and codify the data-sharing and governance process.

Fourth, the Army must protect all forms of HCE data—both personally identifiable information (PII) and protected health information (PHI)—at rest and in transit; doing so ensures that the Army meets legal and regulatory requirements while also maintaining the bond of trust with its personnel. Fifth, individuals with data release authority, and those who conduct analysis or inference based on Army data, must receive appropriate training and certification. Here, the HCBD strategy recognizes the expertise required in the emerging field of data science and calls for the implementation of certification standards for those who manage, analyze, and use HCE data. Sixth, the Army must establish a standardized process for the use of disparate data within the HCBD framework. This principle recognizes there are some data assets that, while unclassified, are highly sensitive and thus great care must be taken with their use and sharing (for example, Provost Marshal General data, security clearance data, and others).

From Strategy to Operationalization

After publishing the HCBD Strategy, the working group moved toward preparing the HCBD implementation plan, which was approved in August 2017. The plan addresses governance, ethical oversight, phasing and tasks for implementation, data management, and associated technical processes needed to support the HCBD enterprise. Perhaps most important, the implementation plan establishes the HCBD Steering Committee as subordinate to the Army Data Board, and consists of senior leaders from the Assistant Secretary of the Army for Manpower and Reserve Affairs, Deputy Chief of Staff G1 (Personnel), and Deputy Chief of Staff G8 (Resources). The committee is tasked with directing the HCBD data governance process across the Army. For example, the plan imbues the committee with the power to review disputed requests for data access and make adjudicative decisions on sharing data. Additionally, the committee shall establish data-sharing criteria; routinely review the ethical, moral, and legal sufficiency of conducting certain high visibility analysis projects; review and manage the HCBD data use agreements; and review enterprise audits.

Beyond governance, the HCBD implementation plan also establishes three categories of analyses supported by HCBD—descriptive statistics, policy analysis, and research. The implementation plan recognizes that several research and analysis organizations already have extensive data access and management policies, such as the Office of Economic Manpower Analysis at West Point, Army Medical Command, Army Research Institute, and others; those organizations may continue to operate as they have, while still leveraging the governance capabilities offered by HCBD. However, many organizations across the Army do not have a long history with data analytics and are not resourced to establish their own information technology and staff infrastructure to support big data

Soldier with U.S. Army Reserve 312th Engineer Company fires M4 rifle paintball gun while opposing force member hides during urban operations training and building clearing procedures, April 18, 2015, at Camp Ripley, Minnesota (U.S. Army/Timothy L. Hale)

analytics. The implementation plan therefore calls for those organizations to turn to the Person-Event Data Environment, HCBD's enterprise architecture for big data analytics and data management.

The PDE

The Person-Event Data Environment, operated by the Army Analytics Group and its Research Facilitation Laboratory, serves as a key enabler for the HCBD's success. Initially established in 2008 as a business intelligence platform with a limited scope focused on civilian personnel forecasting, the PDE added capabilities over time based on emerging project requirements. Below, we describe three watershed events occurring over the last decade that brought the PDE to the forefront of the Army's big data solution set.

First, in 2008, Army senior leadership set a course toward addressing the suicide problem after the Service's suicide rate exceeded the national average for the first time. Because suicide is such a low base-rate event (approximately 30 suicides per 100,000 Soldiers), a significant amount of data from a wide variety of sources was needed to properly study the phenomenon, with even more data needed to build predictive models. What emerged was the Study to Assess Risk and Resilience in Servicemembers (STARRS), an epidemiological and neurobiological study of suicide involving world-renowned scientists.[6] Also in 2009, Army senior leadership directed the creation of a resilience development program known as Comprehensive Soldier Fitness that was designed to address the endemic stressors of Army life.[7] Because resilience is characterized by equifinality (that is, there are many pathways to becoming resilient and resilience is evident in many aspects of a Soldier's life), measuring the program's effectiveness required massive amounts of personnel, deployment, training, family, and medical data, and the data needed to be gathered at multiple points across a long period of time. Lastly, subsequent to an extensive legal review in 2012, Army senior leadership signed an agreement with the University of Pennsylvania to allow a consortium of researchers from across the United States to use the PDE and its data to answer important research questions related to the mental and physical health of Soldiers; though the research is done by consortium researchers, the projects are governed by Army personnel.

In all three cases, each project led to major advancements in the PDE system or changes in philosophical approaches to how the Army used its human capital data. For example, while the STARRS project led to the accumulation of a vast array of data from across DOD to study suicide, this also opened a pathway for repurposing the data for use by other research teams focused on different topics. Here, processes were developed for establishing Data Use Agreements (DUAs), and electronic workflows were developed to speed the review and approval processes that trigger access to the data. The emergence of the Comprehensive Soldier Fitness project focused attention on the need for a comprehensive and integrated human subjects research governance structure within the PDE. Here, every project conducted in the PDE is reviewed by a Human Protection Administrator for compliance with Federal and DOD guidelines related to the ethical and legal protection of human subjects, and, if required, projects undergo additional reviews by external scientists and are later reviewed by Institutional Review Boards. Lastly, the project with the University of Pennsylvania showcased the value of the Army's data for not only answering important Army research questions but also examining important problems affecting the public, such as cardiovascular disease. Here, the Army approves specific research questions posed by the consortium of university scientists. The Army in turn benefits from the knowledge created, but does so at a significant financial discount because the university funds most of the research from an external grant.[8] Arrangements like the one described here not only highlight the value of Army data but also underscore a path toward decreasing the cost of research if handled carefully.

Why Use the PDE?

While the PDE's success in the last decade can be in part attributed to the amassing of data, there is more to the story than the clichéd, "If you build it they will come." In fact, a common question we receive is, "What is PDE's special sauce?" The answer is not an advanced machine-learning algorithm, high-performance computing, or hard-to-get data; while the system has each of those, they are not what set PDE apart. Rather, what makes PDE special is the philosophical approach taken toward data management, which can be summarized in five tenets.

First, privacy concerns coupled with the ethical and legal use of data within the PDE is paramount. Beyond the human research protections governance process previously described, data within the PDE are also carefully managed by coupling security requirements and procedures outlined in the Army chief information officer/G6 ADMP with data de-identification best practices from industry. Typically, all individual identifiers within datasets provided by stewards are either completely removed as the system ingests the data or, in some cases, are encoded with special keys that are kept on separate systems with separate encryption and firewalls. Some "low density" data, such as military or civilian rank, unit identification codes, medical conditions, and others are merged into groups. For example, colonels and general officers are typically grouped into a single "senior leader" group, and exceptions to this policy are granted rarely in order to prevent re-identification based on demographic characteristics. So a study intending to analyze data on female African American general officers who are aviators, flew for the 101st Airborne Division, and later developed a heart condition likely could not be supported by PDE due to data policy restrictions. Data within the PDE are held in separate enclaves—the "Staging Enclave," where data are merged to support the need for each project, and the "Analysis Enclave," where data are first provided to a research team after being encoded for a second time—and researchers are never given access to the Staging Enclave. When data are requested for projects, system administrators monitor automated data transport programs that assemble, de-identify, and position data for the research team. In addition, for sensitive datasets (for example, law enforcement and security clearance information), two-person controlled access is required, and, even then, administrator access is partitioned based on domain (that is, some administrators may access only medical data while others may access only personnel data). Lastly, the PDE is fully auditable, with each action within the system being logged and monitored.

Second, the PDE brings the researcher to the data rather than sending data to the researcher. One does not have to search hard for examples where PII or PHI on DOD personnel was exposed due to losing a laptop or handheld device. Quite simply, the PDE is a private cloud computing environment, and data that support approved projects within the PDE are not allowed to leave the cloud. Thus, researchers may access the PDE cloud from anywhere in the world at any time—even from an aircraft—provided they have a network connection, an approved PDE account, and a computer with a Common Access Card reader and required software. However, they cannot download data to their local machine. When researchers complete their analysis and want to export their findings out of the PDE, privacy specialists review the export request within 24 hours to ensure that item-level data are not part of the export package. This philosophy balances the on-demand and accessibility needs of the consumer with control and protection over DOD data.

Third, while still operating within its governance structure, the PDE will remove barriers to data access and lift administrative and technical burdens from the research and analysis community whenever possible. For example, assume that the Army's Medical Research and Material Command funds 10 projects at 10 universities, with each project requiring data from 10 different DOD systems. Theoretically, this would require the preparation, staffing, and approval of 100 different DUAs between the Army and 10 different universities, each with its own bureaucracies, potentially resulting in thousands of hours spent getting the DUAs approved while losing even more hours *not* doing research. Additionally, all 10 universities would theoretically have to accredit their own systems to meet DOD Information Assurance requirements to store military data, in turn creating additional burdens on the Army staff. Our own internal analyses suggest that researchers working DOD-sponsored data-intensive projects within the human capital domain spend approximately 60 percent of their time and financial

Servicemembers aboard USS *Dwight D. Eisenhower* paint starboard anchor gold, commemorating ship earning Retention Excellence Award for 2016, Norfolk, Virginia, March 28, 2017 (U.S. Navy/Anderson W. Branch)

resources accomplishing these administrative tasks, leaving only 40 percent of the remaining resources to be applied toward doing the actual research; the numbers look even worse when university overhead is factored into the equation. Does it not make sense to do this in an enterprise fashion? Should researchers not spend more time doing what they are trained to do—scientific research—rather than focusing most of their efforts on meeting administrative burdens? The PDE takes an enterprise approach to the data acquisition process such that the PDE data acquisition team writes and staffs "omnibus" DUAs for the PDE system, rather than for individual projects, thus making them scalable. Likewise, researchers using the PDE do not have to be concerned with system accreditation problems because they are simply accessing a secure cloud; the Information Assurance requirements for the PDE are handled by the system administrators.

Fourth, the PDE is a "digital data commons" that should be used by the widest possible audience, necessitating breadth and depth of data and analytics capabilities. Thus, while the data acquisition team typically acquires new data assets when a new "demand signal" emerges (that is, a new project), the team also acquires data in anticipation of future needs. Likewise, the PDE offers a wide variety of statistical platforms, such as R, SAS, SPSS, STAT-A, Mplus, and others, thus precluding individual researchers from having to purchase the software themselves, which further brings down the cost of DOD research. Because the PDE is a scalable architecture that allows for layering in additional technical capabilities, the system can adapt to new technology like the recently acquired Hadoop cluster that allows for massive parallel computational processing, or the inclusion of other high performance and artificial intelligence capabilities being explored now, such as IBM's Watson, Google's DeepMind, C3IoT's Ex Machina, and others.

Fifth, though the data mantra may be "acquire once, share many times," the PDE staff respects the rights and responsibilities of data providers. For example, data within the PDE are categorized in three ways. Data in the least restrictive "open" category are typically DOD assets that are widely shared across many organizations repeatedly (for example, demographic data) and are available to any PDE user without additional reviews by data providers. Requests for data in the "restricted" category triggers the workflow engine to send notifications to data stewards assigned to organizations providing data to the PDE. Once the stewards receive a notification, they may log into PDE, review the research protocols, communicate with the researcher requesting the data, and finally vote for or against access to the data. The final

Commander of Combined Joint Task Force–Operation *Inherent Resolve* and XVIII Airborne Corps reenlists paratroopers of 2nd Battalion, 325th Airborne Infantry Regiment, 2nd Brigade Combat Team, 82nd Airborne Division, near Bartallah, Iraq, February 1, 2017 (U.S. Army/Loni Ayers)

"closed" data category walls off the associated data from all but those researchers who are invited to use it by the provider. Typically, data in the closed category are collected by the researchers themselves in the field and they do not wish to share the data with other researchers until their project is complete. However, in keeping with the digital data commons theme, we highly encourage those controlling closed data to transition them to a less restrictive category once they complete their project.

Research Under Way Now

In its current configuration, the PDE supports approximately 50 research and operational analysis projects annually, and all the projects fall within the human capital domain. As the HCBD program of record emerges, the concept plan calls for investments to scale up both PDE's technical capability and staff to meet the increasing demands of the Army's broader research and analysis community. Although the research domains supported by PDE varies widely, we highlight two projects that are likely of great interest to *JFQ* readers.

The Complex Behavior Models Project

It is clear from the Comprehensive Soldier and Family Fitness training program and the broader Ready and Resilient capabilities that the Army has placed significant emphasis and resources behind a preventive approach toward improving Soldier psychological health and resilience. While developing resilience may be the Army's proximal goal, a distal goal is to improve overall personal readiness of those serving in uniform. And there is little question that improved personal and unit readiness are needed given the recruiting and retention landscape noted by senior Army leaders: Only 400,000 young people become eligible for military service each year, and of those, over 250,000 are needed to meet national recruiting requirements across all Services; within the Army, approximately 20 percent of Soldiers contracted never make it to their first duty station; approximately 40 percent do not complete their first term of enlistment; only approximately 40 percent of West Point and 4-year scholarship ROTC graduates serve past 10 years.[9] When taken together, the annual personnel churn within the U.S. military costs billions of dollars while also degrading military readiness.

While we readily admit that there is no substitute for good leadership and innovative recruitment and retention strategies, the use of predictive analytics should support these strategies. Focusing more narrowly on preventing involuntary attrition and medical readiness of those Soldiers in uniform, the Army Resiliency Directorate launched an initiative 3 years ago known as the Complex Behavior Models (CBM) project that couples advanced machine-learning methodologies with the power of the PDE's data stores. Because Soldiers attrite for many different reasons—personal choice, legal problems, medical ailments, and others—the goal of CBM is identifying health and resilience characteristics of Soldiers that in turn influence personal readiness. To accomplish this, a team of scientists integrated over 40 PDE datasets—and intend to double that number in the coming years—to develop a suite of models that can predict emerging problems, which could result in involuntary attrition or medical non-deployability with a reasonable level of accuracy.

The data requirements for CBM are massive and likely could not easily be managed outside of a system like the PDE. For context, a single integrated CBM dataset focused only on the Active-duty component consists of 387 columns and over 25 million rows of data, resulting in over 9.8 billion cells of data. While this is a lot of data, it is admittedly relatively small when compared to data processed by organizations such as Google, Facebook, Amazon, and others. Yet CBM's goals are to compute outcomes much more complicated than online purchasing decisions or whether someone will "like" another's posting. Here, CBM is using data to understand highly complex behavioral outcomes, and the computational power required to run these analyses both continuously and on demand is significant.

Insider Threat

The insider threat (InT) of malicious behavior by insiders, whether it is on a network or violence within the workplace, continues to be a challenge within DOD. Executive Order 13587, *Structural Reforms to Improve the Security of Classified Networks and the Responsible Sharing and Safeguarding of Classified Information*; the National Insider Threat Policy; DOD Instruction

5205.16, *DOD Insider Threat Program*; and the Army Insider Threat Directive provide guidance on the establishment and conduct of InT programs. We know from the work of the Defense Personnel and Security Research Center and others that the InT problem is complex—there are many reasons why someone decides to commit an InT act, and there are usually many smaller antecedent indicators that typically go unnoticed. Despite the wave of media coverage when they occur, InT acts are exceedingly rare events. Statistically, this rarity makes predicting an InT act extremely challenging.

Taking a cue from private sector companies such as JP Morgan Chase, Goldman Sachs, LexisNexis, and others, scientists working within the PDE are applying its data to machine learning and other statistical methodologies to better understand the InT problem. Statistically, it is much easier to accurately predict a large population of people who probably will *not* commit an InT act than it is to accurately predict specific individuals who probably *will* commit one. Thus, the goals of the Army's InT research and analysis program is to use de-identified data to accurately pool a small population of individuals who, based on their behavioral risk factors, are at higher risk for committing an InT act than those in the large population of those who clearly are not at risk. Though the work is still in the nascent stage, three InT statistical models emerging from the project perform reasonably well, though much work is yet to come.

Challenges Over the Horizon

Though HCBD and its enablers such as PDE represent a significant step forward in helping the Army use its data to better "see itself," there are several challenges that must be addressed in the next few years to ensure that the Army continues to protect its most valuable asset, its people. For example, though emerging technical capabilities within the big data domain suggest that we can create new knowledge, great care must be taken to avoid causing harm. Stated differently, just because we *can* do something with data, the more important question is, "Should we?" The HCBD Steering Committee is charged with addressing this concern. Returning to the CBM and InT examples, once the machine-learning models are validated, how will Army senior leaders decide to operationalize the models within an ethical, legal, and moral governance framework? One option is to transition these models out of the PDE research environment and later integrate them into carefully governed leader decision support tools to assist in making Army-, unit-, and individual-level decisions that should help stave off some of the Army's readiness, attrition, and InT challenges. And though the PDE system administrators go to great lengths to protect anonymity, is there a certain point when so much data are merged that the current PDE data management policies are not sufficient to protect that anonymity? A recently launched project within PDE will run for the life of the system and attempt to answer this question, providing regular recommendations to leadership for policy adjustment. Finally, despite the fact that PDE operates with DOD-standard firewalls and encryption, at what point does the merger of a certain number of unclassified datasets raise the risk to the point where the data should be classified? A working group within the HCBD community is tackling this concern now.

Though what we present here is viewed through a decidedly Army lens, the challenges described are not altogether different than those facing other Services; there are many commonalities. While the Defense Human Resources Activity recently created the Office of People Analytics to provide policy guidance to the Services in the coming years, each Service will likely pursue a human capital data analytics solution set that best meets its needs. For the Army, the HCBD initiative and the PDE both represent an effort toward getting actionable information in the hands of leaders quickly while also protecting the Army community members' privacy. Regardless of each Service's chosen path, the paramount requirement before us all is to create systems that balance the data analytic needs of leaders while strengthening the bond of trust with our Servicemembers. JFQ

Notes

[1] Karl F. Schneider, David S. Lyle, and Francis X. Murphy, "Framing the Big Data Ethics Debate for the Military," *Joint Force Quarterly* 77 (2nd Quarter 2015), 16.

[2] Jody Westby, "The Government Shouldn't Be Lecturing the Private Sector on Cybersecurity," *Forbes*, June 15, 2015, available at <www.forbes.com/sites/jodywestby/2015/06/15/the-government-shouldnt-be-lecturing-the-private-sector-on-cybersecurity/#607d6354621b>.

[3] Michael Zimmer, "Mark Zuckerberg's Theory of Privacy," *Washington Post*, February 3, 2014, available at <www.washingtonpost.com/lifestyle/style/mark-zuckerbergs-theory-of-privacy/2014/02/03/2c1d780a-8cea-11e3-95dd-36ff657a4dae_story.html>.

[4] *Family Educational Rights and Privacy Act*, *U.S. Code* 20 § 1232g; 34 CFR Part 99 (1974).

[5] Department of Defense (DOD) Directive 5400.11, *DOD Privacy Program* (Washington, DC: DOD, September 1, 2011), available at <www.disa.mil/~/media/Files/DISA/About/Privacy-Office/540011p.pdf>.

[6] Anne M. Gadermann et al., "Prevalence of DSM-IV Major Depression among U.S. Military Personnel: Meta-Analysis and Simulation," *Military Medicine* 177, suppl. 8 (2012), 47.

[7] Rhonda Cornum, Thomas D. Vail, and Paul B. Lester, "Resilience: The Result of a Totally Fit Force," *Joint Force Quarterly* 66 (3rd Quarter 2012), 28.

[8] Loryana L. Vie et al., "The Person-Event Data Environment: Leveraging Big Data for Studies of Psychological Strengths in Soldiers," *Frontiers in Psychology* 4, no. 934 (December 2013); Loryana L. Vie et al., "The U.S. Army Person-Event Data Environment: A Military-Civilian Big Data Enterprise," *Big Data* 3, no. 2 (2015), 67–79.

[9] C. Todd Lopez, "To Become 'Force of Future,' Army Must Fix Personnel Churn," *Army.mil*, June 26, 2015, available at <www.army.mil/article/151308>.