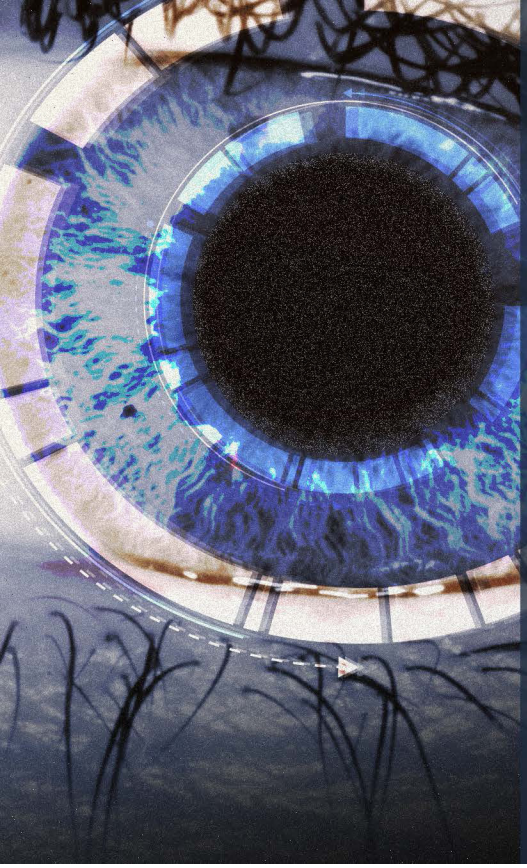

Modernizing Intelligence Operations in Africa: Enhancing the Intelligence Process Through Data Science

by Colonel Chris Tomlinson,
Chief Warrant Officer 3 Felix Rodriguez Faica,
Chief Warrant Officer 2 Ryan Harvey,
and Mr. Keith Hickman

Introduction

The U.S. Army Intelligence and Security Enterprise and other members of the greater intelligence community are not immune from the often-repeated paradigm of rapidly increasing data and emerging technologies producing more information than can be accurately processed and understood. The Department of Defense Data Strategy recognizes the need for a systemic approach to attain analytic maturity to gain information superiority, highlighting the need for "data at speed and scale for operational advantage and increased efficiency."[1] The Army Africa Data Science Center's (ADSC's) application of data science methodologies and technologies has modernized the U.S. Army Southern European Task Force, Africa (SETAF-AF) G-2's ability to analyze and process vast amounts of data. By taking a deliberate, proactive approach to integrating artificial intelligence (AI) and machine learning (ML) and by incorporating data science and engineering (DS&E) to target this information explosion, ADSC provides a problem-solving approach focused on capturing efficiencies in the intelligence process. Using ADSC as a case study, this article illuminates the increasingly pivotal role DS&E plays in enhancing the intelligence warfighting function throughout the U.S. Africa Command (USAFRICOM) area of responsibility (AOR).

ADSC's mission is to provide customized AI and ML capabilities that enable intelligence analysts to answer SETAF-AF and USAFRICOM priority intelligence requirements (PIRs) more efficiently and effectively.[2] This is especially vital in a resource-constrained theater. ADSC accomplishes this in four ways:

- Focusing on improving data literacy across the force, which supports Army Data Plan 2022 and highlights the urgent need for a data-literate workforce.
- Leveraging the geographic expertise of theater-embedded engineers.
- Co-locating DS&E teams directly with intelligence analysts.
- Building modern analytical products and automation on government-furnished cloud technology.[3]

Through a combination of analysis and vignettes, this article highlights what four years of experiential learning have shown: that integrated DS&E teams can have a transformative impact on intelligence operations in the African theater and, by extension, other theaters. These efforts enable intelligence analysts to produce more comprehensive and timely intelligence products, ultimately increasing the commander's decision advantage.

A Foundation of Data Literacy

According to the U.S. Bureau of Labor Statistics, future demand for data-literate workers will increase in every sector of the economy, led by increasing adoption of complex data solutions and infrastructure in fields such as healthcare, finance, transportation, and utilities.[4] These civilian sectors seek to exploit the transformative potential of advanced analytic techniques to achieve better results, including improved patient outcomes, fraud detection, and traffic and safety optimization, while maximizing security to address growing cyber threats.[5] Ensuring a data-literate workforce is essential to achieving these benefits because these industries

will continue to evolve and become more overtly data-centric. Developing this skilled workforce requires an effort from the whole organization. Senior leaders must understand and leverage organizational data capabilities and requirements, analysts must ask more complex questions, and data teams must build solutions that support this model.

These same considerations apply equally in a military context. With the exponential growth of information accessible across all classification levels, military intelligence professionals find it increasingly difficult to triage vast amounts of data to respond promptly to PIRs. Concepts such as pattern recognition, anomaly detection, and predictive modeling are all viable approaches to solving these problems, and they all require a data-literate organization.[6] Nevertheless, there are many pitfalls along the path to organizational analytical maturity. According to the Army Data Plan of 2022:

> *The Army is increasing data literacy across Soldiers and civilians. . . . However, to increase change at scale, the Army needs to increase the basic data skills for generalists that benefit from greater accessibility to quality data to improve daily decisions, that is,* ***citizen analysts benefiting from our data democracy****.*[7]

Data literacy can be acquired in several ways, all entailing individual intellectual curiosity and perseverance. This can be encouraged across the Service through a combination of institutional, operational, and self-development opportunities following the Army training domain framework.[8] ADSC's efforts to improve data literacy within the SETAF-AF intelligence enterprise will fundamentally reshape how intelligence analysts think about the ways data supports the intelligence process. This evolution will improve the quality of intelligence products the intelligence warfighting function provides commanders and staffs to enhance situational awareness and maximize decision space for military operations.

ADSC leads and develops in-person courses and regularly works with analysts and leaders on complex data projects to rapidly improve unit, team, and individual data literacy. Incorporating the recommendations in this article will enable units to field data-centric teams at the appropriate echelon to meet their force data literacy goals.

Data Science and Engineering Structural Best Practices

Incorporating DS&E capabilities into an intelligence organization can increase the efficiency of processes that identify threats, assess risks, and inform decisions in real time while parsing quantities of information that would otherwise be insurmountable. However, achieving these lofty ambitions requires the adoption of specific principles to maximize efficiency and effectiveness, including co-locating developers with analysts and developing regional expertise. Experiential learning with ADSC has identified a "hub-and-spoke" model as the preferred structure to achieve these goals.

Co-location is vital to effective collaboration. A major advantage of the ADSC structure lies in the physical co-location of engineers with analysts. In the private sector, companies allocate billions of dollars annually toward market research to build a deeper understanding of consumer preferences and requirements, thereby maximizing their profit potential by accurately addressing their customers' needs. In software development, this often entails identifying a precise problem (i.e., the consumer requirement) and providing an effective solution that saves time or resources and improves workflow efficiency. Similarly, eliminating the divide between the intelligence analyst (i.e., the consumer) and the data scientist improves the ability to identify, refine, and prioritize requirements while shortening the time needed to develop and implement technical solutions. The Department of Defense and the intelligence community are uniquely situated to position DS&E teams alongside users.

Augmenting the nuanced depth of knowledge provided by intelligence professionals with a niche technical capability allows for quick prototyping of effective and efficient analytic solutions to meet a commander's evolving requirements while eliminating communication barriers. For this reason, DS&E teams need to be managed and staffed at the most tactical echelon possible, with support from and reachback to higher echelons.

Regional expertise generates effective analytic solutions. In addition to the efficiency benefits realized by co-locating DS&E teams with analysts and leaders, regional expertise is critical for effective data solutions. Every command faces a unique challenge presented by its distinctive geography and mission focus that requires time to develop domain understanding and expertise. For instance, some simplified examples of this complex problem include: USAFRICOM units often monitor instability ahead of potential crisis support operations, U.S. European Command and U.S. Indo-Pacific Command units narrowly focus on strategic competition, while U.S. Southern Command units tackle issues like human and drug trafficking. The datasets needed to answer questions based on these discrete missions are often very different, as are the applications built on those datasets. For instance, consider an analytical tool that gives insight into how a commander's PIRs are being answered. Nearly everything in this tool will be different from one command to the next, including the PIR, workflow, information presentation, etc. Additionally, intelligence enterprise datasets are tightly controlled, while operational domain information, such as human resources or logistics, may have different rules and applications.

ADSC acts as an intelligence multiplier by applying data engineering and automation techniques to quickly aggregate and identify valuable information,[9] in a meaningful way for SETAF-AF G-2 personnel. This capability is extremely valuable in a resource-limited and restricted collection environment.

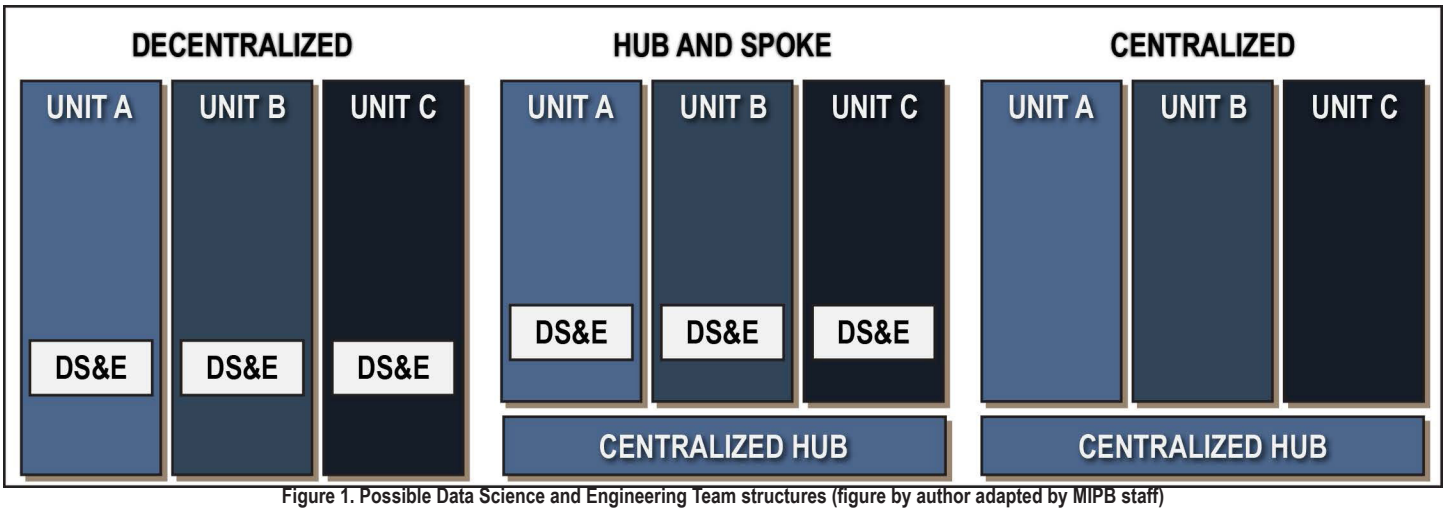

Figure 1. Possible Data Science and Engineering Team structures (figure by author adapted by MIPB staff)

Identified benefits support a hub-and-spoke model. Organizations scaling their data capabilities need technical reachback. Technical support provided at higher echelons substantially reduces the time spent on tasks common to all DS&E teams, including setting up infrastructure, finding development resources, streamlining collaborative projects, and implementing project management practices. This structure is often referred to as a hub-and-spoke model.[10] One hub may serve several embedded teams or "spokes." For instance, a central hub at a joint combatant command might support DS&E teams embedded in several service component commands. A hub might consist of a core group of data engineers and software developers, while a spoke refers to a supported theater-embedded data team such as the ADSC. A hub's primary concern is enabling embedded data teams by providing technical infrastructure and reachback, whereas a spoke directly answers RFIs from units. Compare this model with other configurations (see Figure 1). Centralized DS&E models might hold all data resources in a single space removed from users and, therefore, suffer from a lack of regional expertise. Further, the value proposition of a centralized DS&E team vanishes when considering the need to train, familiarize, and integrate with new commands instead of having organic teams in place. A decentralized model might be efficient but suffers from stovepiping and a lack of central support. ADSC has informally implemented a hub-and-spoke model by building relationships with other command DS&Es, U.S. Army Intelligence and Security Command, and various technical teams across the intelligence community.

The value of data science and engineering with applications and use case. Many companies seek the transformative power of advanced analytic techniques to optimize profits, service levels, and physical or digital products. Even cursory research finds myriad examples of DS&E applications and use cases across every industry sector. Use cases refer to specific data science techniques such as pattern recognition, anomaly detection, and predictive modeling.[11] Example applications include anomaly detection to improve cancer detection methods in healthcare, pattern recognition to detect fraud in finance, and traffic and safety optimization systems for government entities.[12] In addition to increasing efficiency and productivity, these initiatives must also maximize security to address growing cyber threats while incorporating ethical decision-making practices.[13] These and many other applications and use cases apply equally across the intelligence warfighting function. DS&E teams such as ADSC are experts in developing and deploying advanced applications.

Case Study One: Forecasting Violent Extremist Organization Activity in West Africa

One of the main concerns for the SETAF-AF G-2 analysis and control element (ACE) is providing indications and warnings of threats to U.S. forces and equities in Africa. Among the most persistent of these threats are violent extremist organizations (VEOs)—a significant issue across the USAFRICOM AOR, especially in West Africa. Until 2023, the prevailing methodology for conducting indications and warnings assessments was a manual, PowerPoint-based workflow, relying on analysts to interpret large clusters of dots on a map subjectively over long time horizons (see Figure 2 on the next page). Moreover, the inability to achieve the necessary granularity with the available data requires assessing areas prone to VEO presence or expansion at the country or regional level.

Due to the lack of objectivity and granularity within this antiquated methodology, the utility of these assessments was limited. For instance, the efficacy of security cooperation initiatives would noticeably increase if the supporting analysis were to more narrowly identify areas of greatest need, down to specific towns or checkpoints. Likewise, force protection measures can be tailored to a small area if clear trends and reliable forecasts exist for threats in that area.

Approach. In February 2023, senior analysts in the SETAF-AF ACE decided to implement a data-driven analysis of the VEO problem in West Africa. Relying on the unclassified armed conflict location and event data dataset as a suitable proxy for VEO events, ADSC and ACE analysts quantified VEO-related

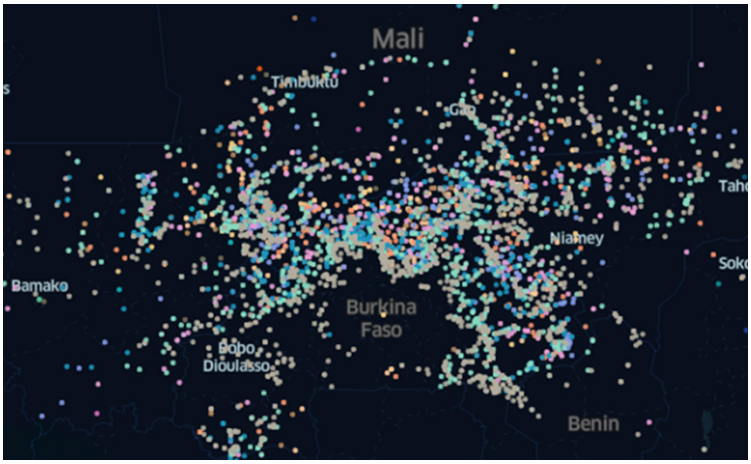

Figure 2. Armed conflict location and event data representation of violent extremist organization activity in West Africa (figure by author)

events to develop narrowly defined geographic forecasts by replicating scientific methods based on AI and ML techniques. Applying the strategies discussed by Andre Python et al. in their 2021 *Science Advances* article, ADSC developed a technique to forecast VEO weekly operational activity by location up to 16 weeks in the future.[14] The underlying location layers are represented by 50-kilometer by 50-kilometer squares published by the Peace Research Institute of Oslo, designed to capture demographic, environmental, and economic information about the squares.[15]

Result. SETAF-AF ACE analysts and ADSC produced a graphical product that forecasts VEO activity at a granular spatio-temporal level to an extent previously impossible with qualitative and subjective methods (see Figure 3). In addition to providing planning and operational support, this product is designed to support a commander's decision-making process for short- and long-term force protection and security cooperation activities. Notably, the product does not replace human-level judgment and only bolsters the qualitative understanding of a given threat assessment.

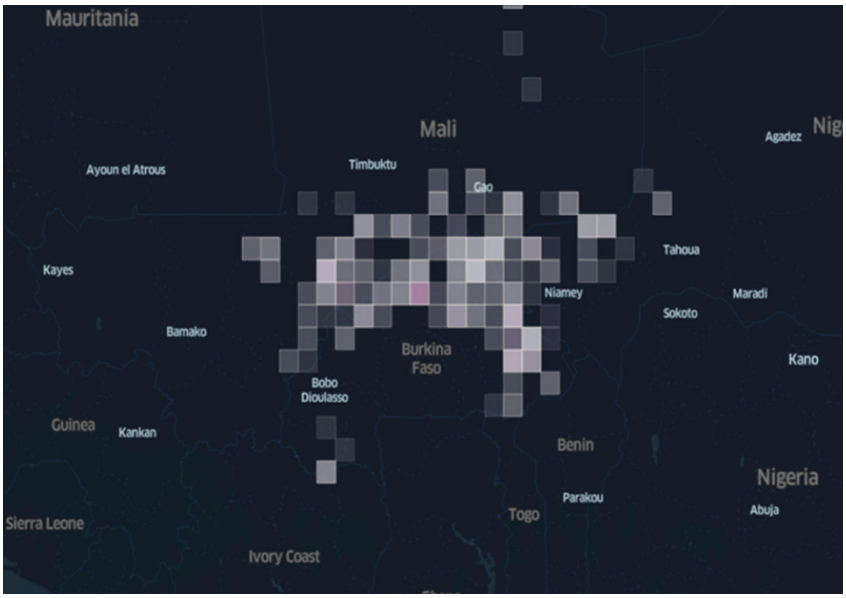

Figure 3. 16-week forecasted violent extremist organization activity on a Peace Research Institute of Oslo grid square (figure by author)

Lesson Learned:

- *Product integration is important.* Data analytics or AI and ML products must be intentionally integrated with existing processes and products, or they will have limited utility and reach. This is vital for new products, such as forecasts.
- *The analyst drives the process.* While the technical product is undoubtedly central to the effort, it can only reflect the analyst's understanding, input, and articulated requirements. Therefore, interaction between the DS&E team and the analyst must happen early and often, which is only possible through co-location, integration, and shared understanding of workflows. If products are intended to support multiple organizations or echelons, stakeholders from all parties should be involved early during requirements generation to maximize applicability.

Case Study Two: Multiple Intelligence Discipline Crisis Support Dashboard

During a recent crisis response operation, SETAF-AF geospatial intelligence analysts monitored hundreds of kilometers of road networks for potential evacuation disruption events, including checkpoints, mobility limitations, and VEO threats.

Approach. After observing several existing workflows, ADSC data scientists embedded with the SETAF-AF analysts identified potential automation projects. ADSC developed Python scripts that emulated keyword and geospatial queries across several intelligence data sources and automatically displayed relevant data in a dashboard. Analysts further requested that the tool provide customized email alerts for all pertinent activity observations. ADSC programmatically overlaid the road networks with a grid system filtering mechanism that displayed activity occurring within five kilometers of areas of interest and provided significant time savings compared to manual monitoring.

Results. This project ultimately achieved three results: cognitive burden shift, error reduction, and information gain. The product shifts the cognitive burden of rote and repetitive data tasks from analysts to computers. The ADSC can write programs that process very large datasets quickly in a meaningful way for ACE analysts, thereby allowing them to focus on critical analysis. Automated scripts like these have the built-in benefit of error reduction because machines process data precisely according to their instructions. Finally, the information gained from seeing many disparate datasets displayed together, such as merging data from multiple sensors or data from multiple intelligence disciplines, is invaluable, though difficult to measure.

Challenges to Adoption

While ADSC continues to demonstrate its value as a force multiplier for the SETAF-AF intelligence enterprise, several constraints have slowed the broader adoption of its technology-based approach and stymied some projects that ADSC has spearheaded.

The greatest challenge comes in educating members of any organization, including analysts and leadership, on the true capabilities that DS&E teams offer and how to use them most effectively. Without education to overcome this challenge, requirements will be either too simple, thus wasting their unique skillset, or too difficult to accomplish, resulting in hours wasted on projects that the team knew would likely never bear fruit. Ultimately, this is not a question of revolutionizing processes. Instead, it is a matter of developing systematic efficiencies that generate results within already adopted practices. Although there has been some resistance to adopting this approach, ADSC offers cutting-edge solutions that can help overcome generational data- and computer-literacy deficits. Within SETAF-AF, increased exposure to ADSC's capabilities across the staff, complemented by a thorough requirements management process, has already started alleviating some of the problems posed by this challenge. Participation in capabilities briefs and support to projects outside the G-2 are good starting points, though support and investment across the command will ultimately be necessary.

Another challenge DS&E entities operating within a Department of Defense construct face is the difficulty of integrating traditional data science tools and platforms into programs of record. These are often specialized commercial off-the-shelf programs with outdated custom modifications and scripting, including security parameters preventing linkages to many external repositories. While acknowledging the legitimate security concerns that inform many of these roadblocks, a commitment to adaptability and modernization through the iterative implementation of DS&E best practices is essential to ensuring the Army retains an advantage over strategic competitors.

Finally, the limited period of analysts' assignment creates an inherent inability to train large numbers of Servicemembers to execute data science tasks proficiently at the unit level. We must therefore rely on contract mechanisms with high costs and uncertain long-term program funding. While this has been a challenge to expanding the ADSC, some benefits will emerge as data scientists continue to deepen their understanding of the problems unique to the USAFRICOM AOR and leverage their depth of knowledge and established automations to present a degree of continuity.

Conclusion

In recent years, incorporating DS&E teams has fundamentally transformed Army intelligence. DS&E encompasses diverse methodologies and technologies to extract valuable insights from vast and varied datasets. Data science has revolutionized how we collect, analyze, and process information by harnessing techniques such as ML, predictive analytics, and geospatial intelligence analysis.

Overall, DS&E teams play a vital role in enhancing intelligence analysis for U.S. Army Soldiers by leveraging advanced analytics, predictive modeling, visualization tools, and automation. By integrating these capabilities into intelligence operations, Soldiers can gain a deeper understanding of the operational environment to produce timely intelligence products that better inform decision makers to achieve mission success.

As the U.S. Army navigates the constantly evolving security landscape of the USAFRICOM AOR, it is imperative to capitalize on the opportunities presented by integrating data science into our established intelligence procedures to stay ahead of emerging threats and challenges. This approach enhances operational effectiveness and increases the efficiency of intelligence procedures. Robust intelligence capabilities remain vitally important in the dynamic and complex operating environment of the USAFRICOM AOR. From countering terrorism and insurgency to addressing regional conflicts and strategic competitors, effective intelligence is paramount for mission success in the USAFRICOM AOR.

Endnotes

1. Department of Defense, *DoD Data Strategy*, October 8, 2020, 2, https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF.

2. Efficiency refers to how quickly a task can be achieved. Effectiveness refers to impact—how much information was gained, how many people were affected, etc.

3. Department of the Army, Office of the Chief Information Officer, *Army Data Plan*, October 11, 2022, https://www.army.mil/e2/downloads/rv7/about/2022_army_data_plan.pdf.

4. "Occupational Outlook Handbook–Data Scientists," U.S. Bureau of Labor Statistics, April 17, 2024, https://www.bls.gov/ooh/math/data-scientists.htm; and Aleksandra Yosifova, "The Best Industries for Data Science Specialists in 2024,"*Career Advice, 365 Data Science* (blog), 11 April 2024, https://365datascience.com/career-advice/the-best-industries-for-data-science-specialists/.

5. Zaid Obermeyer and Ezekiel J. Emanual, "Predicting the Future–Big Data, Machine Learning, and Clinical Medicine," *The New England Journal of Medicine* 375, no. 13 (September 29, 2016): 1216-1219, https://www.nejm.org/doi/full/10.1056/NEJMp1606181; and Xinhu Zheng et al., "Big Data for Social Transportation," *IEEE Transactions on Intelligence Transportation Systems* 17, no. 3 (March 2016):620-630, https://doi.org/10.1109/TITS.2015.2480157.

6. Hsinchun Chen et al., "Crime Data Mining: A General Framework and Some Examples," *Computer* 37, no. 4 (April 2004):50-56, https://doi.org/10.1109/MC.2004.1297301.

7. Department of the Army, Office of the Chief Information Officer, *Army Data Plan*, 3.

8. Department of the Army, Army Regulation 350-1, *Army Training and Leader Development* (Washington, DC: U.S. Government Publishing Office, 1 June 2025), 5.

9. Automation can refer to everything from basic scripting to machine learning applications.

10. Norman Krueger, Tim Gabriel, and Cezar Adam, "The Hub-and-Spoke IT Operating Model: Increasing Innovation and Continuous Improvement," ISG (Information Services Group), n.d., https://isg-one.com/articles/the-hub-and-spoke-it-operating-model.

11. Chen et al., "Crime Data Mining."

12. Obermeyer and Emanual, "Predicting the Future."

13. Zheng et al., "Big Data for Social Transportation."

14. Andre Python et al., "Predicting Non-State Terrorism Worldwide," *Science Advances* 7, no. 31 (July 2021), https://www.science.org/doi/10.1126/sciadv.abg4778.

15. The Peace Research Institute Oslo, https://www.prio.org/.

References

Office of the Chairman of the Joint Chiefs of Staff, Joint Publication 2-0, *Joint Intelligence* (Washington, DC: The Joint Staff, 26 May 2022). Change 1 was issued on 5 July 2024.

"Market Research Industry–Statistics and Facts," Statista, May 16, 2024, https://www.statista.com/topics/1293/market-research/#topicOverview.

COL Chris Tomlinson is the Director of Intelligence, G-2, for the Southern European Task Force, Africa (SETAF-AF) and is operational director of the Africa Data Science Center (ADSC) for SETAF-AF. His prior intelligence assignments include Director of Intelligence, J-2, for the Special Operations Joint Task Force–Operation Inherent Resolve, Deputy Director of Intelligence, J-2, and theater analysis and control element chief U.S. Army Europe. He holds a master's degree in strategic studies from the Marine Corps War College and a bachelor of arts in political science at Texas Tech University.

CW3 Felix Rodriguez Faica is an intelligence planner and common intelligence picture/Army Intelligence Data Platform lead integrator in the intelligence operations division of the SETAF-AF G-2. His previous assignments were at various unit echelons including brigade combat team and military intelligence brigade-theater. He received a bachelor of arts in intelligence studies from American Military University and completed the Digital Intelligence Systems Master Gunner Course.

CW2 Ryan Harvey is an all-source intelligence technician serving as an intelligence planner and performance manager for the ADSC in the intelligence operations division of the SETAF-AF G-2. His previous assignments were at various unit echelons, including brigade combat team and military intelligence brigade-theater. He holds a master's degree in intelligence management from Henley-Putnam University and a bachelor of arts in political science from the University of California, Santa Barbara.

Mr. Keith Hickman is a senior data scientist for the ADSC of the SETAF-AF G-2. He previously served as an Army intelligence officer at a brigade combat team. He holds a juris doctor from Pennsylvania State University and a master's degree in computer science from Indiana University.