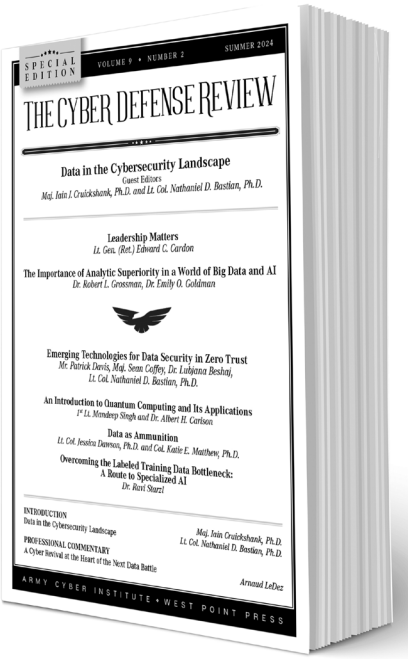
Special Edition on Data in the Cybersecurity Landscape

Major Iain J. Cruickshank, Ph.D.

Lieutenant Colonel
Nathaniel D. Bastian, Ph.D.



We are living through a revolution in human technology driven by digital data. Nearly every piece of technology we use, from cars to refrigerators, now produces—and in some cases, even analyzes—digital data. Furthermore, digital data is now integral to every human endeavor, from dating to warfare. Data is a key component in decision-making and is crucial to warfighting concepts like Information Advantage and Decision Dominance. Thus, it is increasingly important for military and security professionals to understand digital data and the technologies built upon it, such as Artificial Intelligence (AI).

Major Iain J. Cruickshank, Ph.D., is a data and decision scientist (Functional Area 49) in the U.S. Army. He currently serves as a Senior Research Scientist within the Army Cyber Institute at the United States Military Academy (USMA) at West Point and has previous assignments with the 780th Military Intelligence Brigade and the U.S. Army Artificial Intelligence Integration Center. He is also adjunct faculty in Carnegie Mellon's School of Computer Science. Major Cruickshank earned an M.S. in Operations Research from the University of Edinburgh and a Ph.D. in Societal Computing from Carnegie Mellon University. His research interests include development of novel, artificial intelligence based computational social science techniques and multi-modal machine learning, particularly as it applies to military and national security problems.

In an age where digital data reigns supreme, its safeguarding and strategic utilization have become paramount to maintaining effective cybersecurity. This special edition of *The Cyber Defense Review* delves into the multifaceted roles that data plays within the cybersecurity landscape, offering a comprehensive exploration of emerging technologies, innovative frameworks, and strategic approaches to enhance data security and leverage its potential.

We begin our exploration of digital data and cybersecurity in this edition by establishing a new concept: *cyber data sanitization.* "Cyber Data Sanitization: A Cyber Revival at the Heart of the Next Data Battle" highlights the pivotal challenge of data pollution and explores innovative tools and standards for maintaining digital hygiene and optimizing data processing. By repurposing existing cybersecurity tools to clean data beyond malicious activities, cybersecurity expert Arnaud Le Dez proposes a new domain of cyber defense, essential for preserving operational efficiency and supporting AI advancements.

To initiate the research section of this edition, we begin with an article that establishes the importance of data and data-driven technologies through the concept of *analytic superiority*. In "The Importance of Analytic Superiority in a World of Big Data and AI," Drs. Grossman and Goldman outline a comprehensive framework for achieving analytic superiority, which involves collecting requisite data, developing robust analytic models, and deploying these models effectively. Emphasizing the importance of staying ahead in the analytic race, the authors highlight how the U.S. can maintain cyberspace superiority by outpacing adversaries in analytic capabilities.

Lieutenant Colonel Nathaniel D. Bastian, Ph.D., is an Academy Professor and Cyber Warfare Officer at the United States Military Academy (USMA) at West Point. He is also concurrently serving as a Program Manager at the Defense Advanced Research Projects Agency (DARPA). At USMA, Lt. Col. Bastian serves as Division Chief, Data & Decision Sciences, Senior Research Scientist, and Chief Data Scientist at the Army Cyber Institute within the Department of Electrical Engineering and Computer Science. Lt. Col. Bastian earned a M.S. in Econometrics and Operations Research from Maastricht University and a M.Eng. in Industrial Engineering and Ph.D. in Industrial Engineering and Operations Research from Pennsylvania State University. His research interests aim to develop innovative, assured, intelligent, human-aware, data-centric, and decision-driven capabilities for cyberspace operations.

A critical component to unlocking data's value for the warfighter is the security of data. In "Emerging Technologies for Data Security in Zero Trust Environments" Davis et al. evaluate cutting-edge technologies that promise to fortify Zero Trust Architecture (ZTA) data security capabilities. By focusing on differential privacy, confidential computing, homomorphic encryption, blockchain, and alternative computing methods, this article provides a forward-looking perspective on how these innovations can secure data throughout its lifecycle. The authors meticulously assess each technology's potential benefits and drawbacks, underscoring the need for a nuanced approach to integrating these tools into ZTA implementation frameworks.

No discussion about data security would be complete without considering the pending quantum revolution in cybersecurity. "An Introduction to Quantum Computing and Applications" provides a practical guide for understanding and leveraging quantum information science (QIS) technologies. This article elucidates how quantum sensing, networking, communications, and computing can offer strategic advantages in military operations. Emphasizing the importance of informed leadership, First Lieutenant Singh and Dr. Carlson argue that a grasp of QIS applications is crucial for maintaining a competitive edge over our adversaries.

When it comes to defining data and its potential, we should have new mental models for digital data and understand that not all data is good data. Challenging the conventional analogy of *data as oil*, "Data as Ammunition – A New Framework for Information Warfare" presents a compelling case for viewing data through the lens of military ordnance. This paradigm shift offers a more accurate depiction of the strategic

risks associated with data. By categorizing data as ammunition, Lieutenant Colonel Dawson and Colonel Matthew argue, we gain a clearer understanding of its potential threats and how to mitigate them effectively. This framework is particularly relevant in an era where data storage capabilities have skyrocketed, making it crucial to reevaluate our approaches to data security and privacy.

Finally, no discussion of data is complete without talking about the challenges of overcoming data scarcity for data-driven tools, like AI. The scarcity of labeled training data poses a significant hurdle in developing specialized AI for cybersecurity. "Overcoming the Labeled Training Data Bottleneck, A Route to Specialized AI" explores the potential of large language models to generate high-quality, task-specific datasets. By synthesizing existing network intrusion analysis datasets with domain knowledge, Dr. Starzl was able to create a new dataset tailored for zero-day exploit detection. This innovative approach demonstrates the effectiveness of using advanced AI techniques to overcome data scarcity, paving the way for more efficient and effective cybersecurity solutions.

We conclude this edition of *The Cyber Defense Review* with some important reflections on the role of leadership in cybersecurity. "Leadership Matters," by Lieutenant General (Ret.) Edward C. Cardon, reflects on the transformative journey of the U.S. Army Cyber Command. Cardon emphasizes the need for continual adaptation and collaboration among various stakeholders to enhance cybersecurity capabilities. He advocates for leadership to push the envelope in all aspects of cybersecurity, including technologies, concepts, and people because the future is there to be seized by those who constantly hone their skill and maintain the will to execute.

The articles in this special edition collectively underscore the centrality of digital data in contemporary cybersecurity discourse. From the integration of emerging technologies in Zero Trust environments to innovative frameworks for data management and strategic leadership, our contributors offer valuable perspectives on how to harness data's value while safeguarding against its inherent risks. As we navigate the complexities of the digital age, understanding and leveraging data will be crucial to achieving and maintaining cybersecurity resilience. We hope this edition inspires further exploration and innovation in the vital intersection of digital data and cybersecurity.

DISCLAIMER

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