

REPORT DOCUMENTATION PAGE					<i>Form Approved</i> OMB No. 0704-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.						
1. REPORT DATE (DD-MM-YYYY) 04/19/2017		2. REPORT TYPE Master's Thesis			3. DATES COVERED (From - To) AUG 2016 - APR 2017	
4. TITLE AND SUBTITLE Using Big Data to Build a Strategic Warning Process				5a. CONTRACT NUMBER N/A		
				5b. GRANT NUMBER N/A		
				5c. PROGRAM ELEMENT NUMBER N/A		
6. AUTHOR(S) Clarke, Craig S., DOD				5d. PROJECT NUMBER N/A		
				5e. TASK NUMBER N/A		
				5f. WORK UNIT NUMBER N/A		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) USMC Command and Staff College Marine Corps University 2076 South Street Quantico, VA 22134-5068					8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)					10. SPONSOR/MONITOR'S ACRONYM(S)	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S) N/A	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT The U.S. national security apparatus considers the provision of strategic warning to decision makers a critical capability. Using unclassified publicly available information along with the analytics that comprises "Big Data," the United States Intelligence Community can build an intelligence process to inform strategic warning identifying innovative and insightful trends on the future operating environment. The explosion of publicly available information creates the dynamic of "Big Data." Building the appropriate analytical framework along with incorporating the innovative use of Big Data analytics that uses publicly available information as the collection source can establish a strategic warning process.						
15. SUBJECT TERMS Big Data; OSINT; Open Source Intelligence; Strategic Warning; data science; data scientist; correlation; causation; crisis mapping; crowd sourcing; DARPA; National Intelligence Community; Walmart; Amazon; predictive analytics; predictive intelligence; anticipatory intelligence; crowd seeding; algorithm						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT	b. ABSTRACT	c. THIS PAGE			USMC Command and Staff College	
Unclass	Unclass	Unclass	UU	44	19b. TELEPHONE NUMBER (Include area code) (703) 784-3330 (Admin Office)	

United States Marine Corps
Command and Staff College
Marine Corps University
2076 South Street
Marine Corps Combat Development Command
Quantico, Virginia 22134-5068

MASTER OF MILITARY STUDIES

TITLE:

Using Big Data to Build a Strategic Warning Process

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF MILITARY STUDIES

AUTHOR:

Craig S. Clarke

AY 16-17

Mentor and Oral Defense Committee Member: Benjamin Jensen, PhD

Approved: [Signature]

Date: 19 APR 17

Oral Defense Committee Member: Paul Chase Piro

Approved: [Signature]

Date: 19 APR 17

Kevin D. Glatthorn
KDG/asmc 19 Apr 2017

Executive Summary

Title: Using Big Data to Build a Strategic Warning Process

Author: Craig S. Clarke, Department of the Navy

Thesis: Using unclassified publically available information along with the analytics that comprises “Big Data,” the United States Intelligence Community can build an intelligence process to inform strategic warning identifying innovative and insightful trends on the future operating environment.

Discussion: The U.S. national security apparatus considers the provision of strategic warning to decision makers a critical capability. Beginning with the National Security Act of 1947, the IC established a core capability of clandestine intelligence collection, which then informed strategic warning. Despite building an unmatched intelligence collection capability, the end result of the process struggled to inform strategic warning as the IC experienced several failures, Korea in 1950, the rise of Castro, the fall of the Shah in 1979, and the attacks of 9/11 to name a few. Strategic warning fails in these cases and remains elusive because it lacks the proper analytical framework along with clandestine intelligence collection falling short as the appropriate information source. Several initiatives over the years have attempted to reform the IC in reaction to perceived strategic warning failures, including the Intelligence Reform and Terrorism Prevention Act (IRTPA) of 2004.

IRTPA establishes OSINT as an intelligence discipline, which points to publically available information as a collection asset. The explosion of publically available information creates the dynamic of “Big Data.” A widely accepted industry definition articulates Big Data as “high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization.”¹ A review of private industry using Big Data to build strategic warning to gain a competitive advantage provides a model of building a process to inform predictive intelligence. Additionally, the non-profit sector and non-governmental organizations harness the power of Big Data to penetrate areas of ungoverned spaces. Using innovative methods such as crowdsourcing, crisis mapping, and crowd seeding, these organizations tap the power of personal communication technology available to citizens who live in constrained environments enabling collection of their user-generated data. Unstructured information collected from social media platforms informs strategic warning for areas of limited statehood. Big Data solutions are under exploration by the federal government, namely the Defense Advanced Research Projects Agency (DARPA). Innovating with an open source coding approach, the DARPA program focuses on data science, but lacks analytical approaches employed by the IC.

Conclusion: Building the appropriate analytical framework along with incorporating the innovative use of Big Data analytics that uses publically available information as the collection source can establish a strategic warning process.

DISCLAIMER

THE OPINIONS AND CONCLUSIONS EXPRESSED HEREIN ARE THOSE OF THE INDIVIDUAL STUDENT AUTHOR AND DO NOT NECESSARILY REPRESENT THE VIEWS OF EITHER THE MARINE CORPS COMMAND AND STAFF COLLEGE OR ANY OTHER GOVERNMENTAL AGENCY. REFERENCES TO THIS STUDY SHOULD INCLUDE THE FOREGOING STATEMENT.

QUOTATION FROM, ABSTRACTION FROM, OR REPRODUCTION OF ALL OR ANY PART OF THIS DOCUMENT IS PERMITTED PROVIDED PROPER ACKNOWLEDGEMENT IS MADE.

Table of Contents

	Page
Executive Summary	2
Disclaimer	3
Preface.....	5
Introduction.....	6
What is Strategic Warning?	8
The Utilization of Strategic Warning and Why It Struggles.....	10
Publically Available Information, Open Source Intelligence, and Big Data	18
How does Private Industry use Big Data?	21
How does the Non-Profit Sector use Big Data	24
Proposed Courses of Action to Build a Strategic Warning Process	27
Build Upon DARPA's Big Data Program	28
Where Should a Big Data Strategic Warning Process Reside in the IC?	32
Create a Big Data Analytical Framework.....	33
Conclusion and Recommendations for Future Research	35
Endnotes.....	37
Bibliography	41

Preface

An intelligence analyst openly discussing intelligence support to a recent operation is not a typical situation. In 2010, I openly discussed intelligence support in an unusual situation, a documentary for the Public Broadcasting System. Penn State University Television produced *The Geospatial Revolution*, a documentary about the rise of location-based technologies and digital mapping. Wireless technology and smartphones practically cover the globe and allow “digital citizens” to produce personal information with a geo-location. The producers of the documentary interviewed me and the S-2 of the 22nd Marine Expeditionary Unit to discuss Operation Unified Response, the humanitarian response to the 2010 Haiti Earthquake. The discussion focused on how we used crowdsourcing and crisis mapping to direct aid to the places and people in Haiti most in need. The paradox of using geo-located information created by citizens of a failed state in the midst of a natural disaster with personal communication technology posting to social media astounded all of us. Since building a process for operational warning using OSINT, the possibilities of what is achievable with open information continue to fascinate. The seven years since the Haiti earthquake is an eternity in the technological world.

I would like to acknowledge the help of several people in producing this thesis. Thanks to my wife Brianne for her loving support and encouragement. Thanks to Dr. Benjamin Jensen for his mentorship through the process and to the men and women of the Gray Research Center for invaluable research and writing assistance. Thanks to my friends and colleagues who provided me invaluable thoughts not only on this project, but professional guidance throughout my career: Major Chris Mercer, USMC (ret), Jill Clark-Foulkes, Robert Silagi, Phil Autry, Amanda Buck-Ortega, GySgt Carlos Raices, GySgt Colin Crane, and all the men and women of the OSINT community from across the Intelligence Community. Unfortunately, space does not allow for mention of the many Marines and analysts that have influenced me professionally.

Introduction

When an analyst working in the U.S. Intelligence Community (IC) sat down at their desk in 1990, a scarce amount of information powered their analysis as compared to today. Classified intelligence collected through clandestine methods comprised virtually all of the information available for analysis, as the thought of tapping into openly available information to answer classified intelligence requirements remained almost unthinkable. Twenty years later in 2010, the chief of the U.S. Air Force intelligence enterprise declared a reversal of the situation for analysts. “We’re going to find ourselves in the not too distant future swimming in sensors and drowning in data,” said Lt. Gen. David A. Deptula, Air Force deputy chief of staff for intelligence, surveillance and reconnaissance in January 2010.²

Over seven years later, Lt. Gen. Deptula’s prediction on the explosion of information is proving accurate. The amount of openly available information the world produces for collection is staggering. Every 48 hours, the world produces as much information as it did from the dawn of history through 2003.³ This information explosion creates two new dynamics for the IC: a new intelligence discipline of Open Source Intelligence (OSINT) that in turn gives birth to the age of “Big Data.” OSINT is publicly available information (PAI) that anyone can lawfully obtain by purchase, request, or observation that answers a specific intelligence requirement.⁴

The enormity of PAI available for collection creates the age of Big Data. Considering that a definition is at times elusive, what exactly is Big Data? While there are several competing definitions, a RAND study commissioned by the Navy offers the strongest departure point to understand the emerging dynamic:

“...big data is a data set so vast that it stresses the limits of traditional (i.e., relational) databases along four parameters:

- volume of data
- variety of formats, sources, and types
- velocity of searches and data retrieval
- veracity of conclusions based on data.”⁵

The thought of answering intelligence requirements with PAI along with the dynamic of Big Data upends the conventional assumptions of the IC regarding intelligence tradecraft. The modern Intelligence Community as articulated in the National Security Act of 1947 established the IC with the mission to conduct clandestine activities to build strategic warning capabilities to keep the nation safe from future surprise attacks such as Pearl Harbor.⁶ The competing dynamics of OSINT creating Big Data versus clandestine intelligence collection informing strategic warning presents an intriguing question. Can Big Data create a strategic warning system that is more effective than the conventional mode of using classified intelligence?

Technological innovations utilized by private industry successfully creating effective strategic warning provide a model for the IC to utilize Big Data analytics. The Intelligence Community capitalizes on the emerging tech innovations by recognizing the public is now the persistent sensor thanks to wireless technology and social media. There must be a shift from realizing where all previous collection sensors from orbiting satellites to human beings on clandestine missions were all created, tasked, and controlled. This transference results in the tasking of a process, such as crowdsourcing or big data analytics, instead of the tasking of an object, whether human (HUMINT) or electronic (SIGINT, GEOINT).

If the IC is going to match pace with these developing technologies of Big Data analytics to exploit fully the deluge of PAI for driving operational posture, then the Intelligence Community must develop Big Data into an intelligence process. Such a process would not only develop analytics for short and mid-range intelligence forecasts, but also provide predictive long-range trends about the future operating environment. This process would provide strategic warning to policy officials and military leaders on the character and level of security threats that may compel planning and readiness reevaluation.

This paper investigates if an intelligence process for Big Data that would provide strategic warning is possible. First, a literature review establishes the definition of strategic warning and how the Intelligence Community historically has utilized it. Next, Open Source Intelligence and publically available information (PAI) are defined in order to show how PAI creates the dynamic of Big Data. An applied literature review surveying best practices for Big Data from industry and the non-profit sector demonstrates how they gain a competitive advantage and an approach that could enrich strategic warning in the IC. Lastly, suggested courses of action present how to develop a strategic warning process using PAI and Big Data for the Department of Defense. This study concludes with recommendations for future research.

What is Strategic Warning?

A review of U.S. government strategies related to intelligence and national security demonstrate providing strategic warning is a critical capability, yet when articulated the concept remains vague. *The National Intelligence Strategy of the United States of America* currently states that the IC will seek to improve its ability to “convey early warning to national security customers to provide them with the best possible opportunity for action.”⁷ In 2016, the Defense Intelligence Activity issued a strategy that describes achieving an innovative culture with one of the intermediate objectives of enhancing strategic warning by providing insight into adversary’s future intent.⁸ *The U.S. Department of Justice Strategic Plan 2014-2018* states it “will strive to combat espionage by becoming more proactive and preventive in approach, identifying early indications and warnings of traditional espionage and other less traditional espionage-related activities.”⁹ The latest edition of the *National Security Strategy* asserts “enabling indications, warning, and attributions of attacks” develops the ability to deter and defeat attacks from adversaries.¹⁰ The various strategies of the U.S. national security apparatus emphasize strategic warning, but none of them provides a specific definition with details on how to implement.

So, what exactly is strategic warning? Cynthia Grabo's seminal work traces the development of warning intelligence in the IC from the early days of the Cold War forward into the beginning of the Global War on Terror post September 11th. In her work, Grabo defines strategic warning as indications that a future hostile event will transpire against the U.S. or its allies by a hostile state or actor, a development or conflict that affects U.S. security interests, a significant conflict between nations or groups not allied with the U.S., or the threat of a terrorist attack. Grabo concedes that strategic warning does not have a widely accepted definition across national security circles.¹¹

Yet, providing warning is at the heart of the intelligence profession serving as the impetus of creating the Central Intelligence Agency with the National Security Act of 1947, citing the prevention of another Pearl Harbor type surprise. A 1992 study of strategic warning conducted by the CIA argues that "providing policymakers with persuasive and timely intelligence warning is the most important service the Intelligence Community can perform for the security of the United States." The CIA study succinctly defines warning as "a process of communicating judgments about threats to US security or policy interests to decision makers" and is predictive and anticipatory.¹²

The definition of strategic warning seems simplistic and obvious - a forward looking assessment that anticipates threats. Although, as a study on how to improve analysis by The Center for the Study of Intelligence argues the production of strategic warning must be for a customer in the context of a complicated global environment tailored to their specific mission. The detection of activities that are considered dangerous, such as a potential hostile nation massing forces, is often the default view of providing warning. In reality, true strategic warning will assist decision makers by recognizing activities that comprise a forming situation whose consequences could in the future make a significant impact. When preformed correctly, strategic

warning provides leaders with a range of possible courses of action. Strategic warning anticipates a future environment recognizing possible threats and provides space to react with the correct course of action, even possibly preventing the looming threat from forming.¹³ Analysis is the key function in strategic warning in recognizing events that will have a future impact that presently may appear benign and then present the findings in the proper context to a leader.

Grabo lays out in detail how strategic warning is performed as an intelligence function, demonstrating that the analytical rigor is a necessity. The most effective warning analysis is not produced by a thorough and careful examination of current intelligence on a particular topic or geographic region. Rather, the most effective warning intelligence is the product of an iterative process that thoroughly exams all relevant intelligence stretching back weeks to years. Grabo argues this ongoing and constant review of compiled intelligence will produce “a solid basic understanding of the potential enemy’s objectives, doctrines, practices, and organization” enabling an analyst to recognize indications that warn of potential dangerous events. Ultimately, production of strategic warning is neither compiling exhaustive information nor reaching consensus in an effort to predict with absolute certainty, but is assessing all possible outcomes and reaching an analytical judgment on each outcome for a decision maker.¹⁴ An analyst providing leaders future options is the centerpiece in the strategic warning process.

The Utilization of Strategic Warning and Why It Struggles

The historical use of strategic warning within the IC has evolved over time. Often, the emphasis on strategic warning within the intelligence profession ebbs and flows in response to perceived intelligence failures. Providing strategic warning established the Intelligence Community in 1947 in response to the belief that appropriate intelligence analysis prevents Pearl Harbor. In the immediate aftermath of World War II, intelligence leaders and the Truman administration believed strategic warning and clandestine activities worked in unison. The

National Security Council defined the early days of the Intelligence Community with directives called NSCIDs, which stood for National Security Council Intelligence Directives. The first six NSCIDs issued on December 12, 1947 focused the mission of the IC on two main areas, the coordination of clandestine activities and strategic warning.¹⁵ From the beginning, part 6 of NSCID #1 directs the DCI to coordinate the various intelligence functions to produce strategic warning.¹⁶ NSCID #5 prescribes the conduct of espionage and counterespionage operations.¹⁷

Into the early days of the Cold War, the process of collection from clandestine activities informing strategic warning created intelligence estimates on Soviet activities. Early HUMINT activities conducted by military attaches in the late 1940s attempted to provide intelligence estimates on Soviet military intentions for investments in future equipment development. Public displays of their latest hardware in celebrations such as May Day and Revolution Day as well as Soviet Day airshows became targets for collection by attaches armed with clandestine photographic gear. For example, military attaché USAF Major Edison Walters attended the Soviet Day airshow in July 1947 and collected photographic intelligence on aircraft, even providing a HUMINT report on the appearance of Stalin. However, U.S. leaders came to realize these early collection activities provided limited capabilities to inform meaningful strategic warning. Air Force leadership estimated in late 1948 that six to nine months transpired between launching a new Soviet military asset to detection by U.S. intelligence.¹⁸

As Soviet influence grew in Europe and Asia, the IC struggled with not only collecting intelligence, but also in arriving to correct analytic conclusions to inform strategic warning. These shortcomings became painfully apparent with a major intelligence failure in June 1950 - failure to provide strategic warning of the North Korean invasion of the South. General Douglas MacArthur in his role as Far East Commander established in Seoul a secret intelligence office, the Korean Liaison Office (KLO), with the mission to collect on Communist intentions on the

peninsula and provide strategic warning. The KLO covered the activities of communist guerillas effectively with agents operating in both the North and South. Also, the U.S. Army established a secret intelligence operation in Manchuria close to the Korea border that trained ethnically Korean POWs from Japanese World War II units to infiltrate units in the North. These collection efforts informed intelligence reports in 1948 that the North could possess the ability to attack the South as early as 1949, citing logistical improvements and troop buildups.¹⁹

While these efforts continued to effectively report on North Korean activity, the strategic analysis provided to the Truman administration argued that North Korea operated as a “puppet state” of the Soviet Union. These reports established the widely held belief in Washington all the way up to the eve of the June 1950 invasion that the Soviets dictated North Korean actions. Furthermore, U.S. leaders believed the Soviets dictated all Communist activities across their sphere of influence in Asia and Europe. The strategic warning assessments concluded North Korean offensive operations only commenced with a signal from Moscow. North Korea launching an attack surprised Washington. Despite effective intelligence collection, proper analytic judgment failed to inform strategic warning.²⁰ The intelligence failure of Korea proves Grabo’s major argument regarding strategic warning - it relies on the analyst assessing all possible outcomes and providing the decision maker judgments backed by strong analytical rigor. Analysis postulated only one outcome; no attack until the Soviets gave the green light.

In the wake of the Korean intelligence failure, the White House and the Intelligence Community went to work with reforms to make the changes needed in order to be able to produce strategic warning. The reforms targeted the two areas of transforming analysis and improving collection. President Truman appointed Lt. Gen. Walter Smith to take the helm of DCI with the mandate to reorganize the CIA and force the Intelligence Community to work collaboratively in unison. As Grabo would later suggest, DCI Smith delineated a clear division of

analysts working on current intelligence from those working on future estimates by establishing an office that produced the President's daily bulletin and a separate Office of National Estimates to produce strategic forecasting. Also, Smith pushed the other IC organizations performing intelligence functions to work together so that CIA products included their reporting to avoid duplication and strengthen analysis. Not all functioned perfectly during Smith's tenure as DCI, but he is credited with forming the various agencies across the functions of government into something that began to appear as a "community" in order to improve reporting and reformed CIA structure to better produce strategic warning.²¹ In addition, President Eisenhower established on the National Security Council a subcommittee to inform strategic warning of Soviet activities comprised of the DCI and the Chairman of the Joint Chiefs of Staff including several other principals that had purview over the intelligence activities of the U.S. national security apparatus.²²

Following analytical and organizational reforms, collection capabilities received investments reflecting the belief that clandestine activities producing strategic warning are the foundation of intelligence tradecraft. Collection received improvement with developing stronger methods of human intelligence and deep investments into technical collection. Driving HUMINT improvements, Eisenhower directed famed aviator LtGen James Doolittle to study the efficiency of the CIA's ability to develop sources for collection. The Doolittle report of 1954 entitled "The Report on the Covert Activities of the Central Intelligence Agency" is now declassified yet remains redacted in key portions. Doolittle urges the United States to push into new areas of intelligence activities, which may cause discomfort for leaders and even the American people, in order to produce effective strategic warning. "We must develop effective espionage and counterespionage services and must learn to subvert, sabotage and destroy our enemies by more clever, more sophisticated and more effective methods than those used against us. It may become

necessary that the American people be made acquainted with, understand and support this fundamentally repugnant philosophy.”²³ The culture of the U.S. national security apparatus accepts subversion and deception as an acceptable means to improve HUMINT collection to produce strategic warning. This era of improving HUMINT sources and methods is referred to as “the collection revolution” as the ability to collect via clandestine methods vastly improved.²⁴

Technology improvements were the second area of investment to strengthen clandestine collection. President Eisenhower as a commander during World War II developed an affinity for imagery intelligence with its ability to provide a literal photograph of enemy positions and strength. Seeing a gap in clandestine collection via technological means as President, Eisenhower pushed the IC to build IMINT capabilities. Out of this IMINT renaissance of the 1950s, the development of the U-2 spy plane revolutionized overhead imagery collection. Rushed into production by Lockheed in 1955, the U-2 provided the ability to fly reconnaissance mission at extremely high altitudes deep into Soviet territory producing a remarkable imagery collection capability. Eisenhower and the NSC approved every U-2 flight into the Soviet Union and personally reviewed each sortie’s collection. In addition to the U-2, early investments into developing spy satellites for IMINT began to take place at the same time. The Kennedy administration exploited outer space in the coming years, but building a satellite network began this during “the collection revolution.” At the turn of the decade into the 1960s, the U.S. capacity to exploit IMINT for strategic warning remained unmatched by adversaries.²⁵

Even with establishing during the Cold War the foremost intelligence gathering apparatus, the Intelligence Community moved tediously to transform clandestine collection into analysis producing strategic warning. The United States would soon be taken by surprise with events in Cuba with Castro rising along with the Cuban Missile Crisis. Strategic warning mistakes would continue to follow with the Soviet invasion of Czechoslovakia in 1968, the Yom

Kippur War in 1973, the Iranian Revolution in 1979 and Iraq's invasion of Kuwait in 1990; to name a few. Roberta Wohlstetter in her seminal work analyzing why strategic warning falls short argues that failure is rooted in the inability to recognize true signals embedded within the noise of all the collected information. A clear signal is never apparent, but rather several ambiguous ones need distinguishing within the vast compilation of irrelevant information using sophisticated analysis. And in the labor to discern truth from noise, there are two factors against analysts. One factor is timing, delivering the correct signals of the coming event in a timely manner to the appropriate decision makers. The second factor is to the inability of decision makers being willing to contradict accepted strategic assumptions that had become standing beliefs. Wohlstetter argues the remedy is analysis more rigorous and sophisticated coupled with clearly articulated analytical frameworks with flexibility unhindered by conventional mindsets.²⁶

Wohlstetter analyzes the two events, Pearl Harbor and the Cuban Missile Crisis, in order to determine why strategic warning failed in both despite effective intelligence collection efforts. In both cases, the collection disciplines of SIGINT and HUMINT provided a vast amount of information on Soviet intentions in Cuba and Japanese motives in the Pacific. Standing assumptions obscured the discernment of visible signals and created biases causing a decision maker to arrive to a predetermined conclusion. The Kennedy administration working to ease tensions with the Soviets, looking towards a test ban treaty, overlooked potential aggressive behavior. In addition, technological assumptions overestimated the minimum time to deploy an operational ballistic missile. With Pearl Harbor, Washington strongly desired for World War II to remain a one theater war in Europe. Similarly, assumptions regarding the Japanese military believed it lacked the ability to launch a military operation on the scale and effectiveness of December 7, 1941. In the cases of Pearl Harbor and Cuba, small clues pointing to the truth were easier to overlook than countering assumptions and desired outcomes.²⁷

Strategic warning continues to evolve to the present day. To solve the issues intrinsic to collection and analysis, two organizations created with strategic warning in their mission set to deliver predictive analysis increased the size of the Intelligence Community. In 1961, President Kennedy and Secretary of Defense Robert McNamara established the Defense Intelligence Agency in an effort to coordinate the intelligence activities of all the military branches. One of the core functions of DIA is providing intelligence estimates for strategic warning to defense leaders as articulated in the DOD directive that established the organization.²⁸ DIA features strategic warning in its support to senior military leaders to this day. DIA's threat assessments focus on specific countries or non-state actors where conflict could call for a U.S. military response.²⁹ The majority of DIA's strategic warning products analyze possible conflicts within the next five years with an indications and warning analytical template.³⁰

Today at the national level, the Intelligence Community's featured analytical endeavor producing strategic warning is the National Intelligence Council (NIC). Evolving out of several previous attempts to provide intelligence estimates, the NIC directly supports the Director of National Intelligence as the IC's center for the production of predictive analysis and long-term strategic intelligence. The Intelligence Community considers the NIC's strategic analysis the most authoritative analytical assessments available. As is written in the U.S. Code establishing the NIC in 1979, the council is "senior analysts within the intelligence community and substantive experts from the public and private sector, who shall be appointed by, report to, and serve at the pleasure" of the Director of National Intelligence.³¹ Often on assignment from other intelligence agencies to the NIC, these senior analysts are known as National Intelligence Officers (NIO) and currently cover 14 functional and geographic areas as resident experts. The customers of the NIC include the National Security Council, Congress, and the Pentagon. Along with answering national level requirements of these customers, National Intelligence Estimates

are the NIC's marquee product, an authoritative assessment covering any one of the 14 areas providing predictive analysis and strategic warning.³²

Although the analytical tradecraft producing their work is mostly traditional, the NIC takes a decidedly non-traditional approach with personnel. As prescribed in the law, the NIC includes members from outside of the IC and government, including academics and policy experts from think tanks. The goal of the diversity of viewpoints is producing intelligence estimates unbound by institutional groupthink or assumptions without regard to current U.S. policy positions. Unique in their line of products is *Global Trends*, an unclassified consideration of the possible geopolitical and cultural outcomes the future may hold. *Global Trends* is constructed with input from experts around the globe with the intent of widening the analytical perspective and breaking common assumptions held in the U.S. worldview. Even with their innovative approach, Congress sharply criticized the NIC in several instances, especially for failing to provide an estimate on the Al Qaeda before 9/11 and miscalculating Iraq's weapons of mass destruction capabilities in 2002. Although when polled by the Congressional Research Service, most members believe their work benefits from the analysis of the NIC.³³

This literature review of strategic warning reveals the key components comprising a critical capability that provides indications and warnings of threats. A strategic warning process anticipates future threats using predictive analysis to identify present activities, which may appear benign, yet later progression will affect the United States negatively. Intelligence analysis provides the cornerstone to strategic warning assessing a range of possible outcomes rather than attempting to predict a certain outcome, which provides options for a decision maker. Strategic warning struggles when standing assumptions obscure visible signals within the collected intelligence. Failing to challenge assumptions prevents assessment of all possible outcomes. Big Data provides strength to the key components of the strategic warning process in several ways.

An iterative process of thorough examination of all relevant intelligence over time produces the most effective warning. Big Data provides a collection platform of immense depth and breadth providing an analyst the opportunity to go much deeper than when using only classified intelligence. An iterative strategic warning process using Big Data produces the ability to assess all outcomes across the spectrum of possibilities, and especially the opportunity to recognize where standing assumptions obscure signals pointing to unforeseen threats.

Publically Available Information, Open Source Intelligence, and Big Data

As outlined, the Intelligence Community collecting intelligence through clandestine activities to provide warning is a founding principle. As such, the exploitation and analysis of espionage cloisters the work away at different levels of classification, available only to senior government officials or cleared intelligence professionals. Throughout the 20th century, clandestine activities collected the most powerful information available enabling intelligence analysis. The collection with subsequent analysis of classified intelligence is the process that builds strategic warning. A closed source produced the strongest intelligence throughout the 20th century. The IC exerts effort to obtain this information through clandestine methods from a closed source.³⁴

Yet, this dynamic flipped as the digital age of the 21st century transformed the process that creates information. The vast majority of information is now created openly, including the strongest information that answers our strategic questions. Every 24 hours, humans create 2.5 quintillion bytes of new data. This exponential growth means the information created within the last 24 months comprises 90% of data available in the world. A myriad of sources creates new information; sensors collecting environmental data, individuals posting to social media, digital technology creating videos and photos, and commerce generating business, to name only a few.³⁵ With mobile technology, humans are persistent sensors generating information about the world

around them. Currently, 77% of Americans own smartphones with the ability to create videos and photographs.³⁶ Globally, a mobile industry study estimates that by 2020 smartphone owners will double across the world from their current figure of 3 billion to 6 billion with 70% of the population using smartphones. Emerging markets of Africa, the Middle East, and the Asia Pacific will produce close to 80% of these new smartphone owners by 2020.³⁷ Strategic warning requirements focus prominently on these three geopolitical areas along with Russia, making this growth statistic rather notable.³⁸ These emerging markets mean the already staggering amount of data created in geographic areas of interest will continue to grow exponentially.³⁹ In constrained environments, innovative tools even allow for the networking of cellphone and computers at scale without an Internet connection creating two-way information, posted in the future to social media once a connection is established.⁴⁰

Prior to this digital technology boom, only individuals with the necessary means controlled the message of societal discourse as their elite privilege provided access to a media resource which produced the ability to control public societal discourse or create closed information not shared publically. Smart phones with applications developed for streamlined connection provide capabilities for instant information creation and a global audience. Bennett demonstrates the digital age allows individuals to become “the center of their own universes” and “those universes can be very large thanks to the social networking potential of ubiquitous communication technologies.”⁴¹ The connected citizen is now the persistent sensor.

With this transformative digital information process at work, Open Source Intelligence (OSINT) emerges as the newest discipline in intelligence. The same law that reformed the Intelligence Community creating the DNI, the Intelligence Reform and Terrorism Prevention Act of 2004, establishes OSINT as an intelligence discipline, mandating the establishment of the Open Source Center.⁴² The National Defense Authorization Act of 2006 defines OSINT:

“publicly available information (PAI) that anyone can lawfully obtain by purchase, request, or observation that answers a specific intelligence requirement.”⁴³ PAI includes not only user-generated social media content, but also the deep web, the dark web, library databases, commercially available imagery, technical data and publication, gray literature, and more. Information becomes intelligence when it answers a requirement, thus PAI becomes OSINT when open source collection answers a requirement.⁴⁴

While also establishing the intelligence discipline of OSINT, the explosion of PAI creates the dynamic of “Big Data” for the Intelligence Community. Intelligence analysts are “swimming in sensors and drowning in data” as collection efforts amass prodigious amounts of PAI while the tradecraft to engage big data in order to make sense of it has not developed.⁴⁵ As an intelligence collection capability, publically available Big Data is the most easily attained. Technological innovations demonstrate that the public is the persistent sensor with the strongest information available from sourcing PAI. This notion is where the culture of the IC must shift from realizing all previous clandestine collection efforts from IMINT to HUMINT to SIGINT were created and tasked under government control. This transference results in the exploitation of the Big Data process, which is not under government control and necessitates becoming comfortable with the notion of solving our hardest intelligence problems with collection that is no longer clandestine.

The key components comprising OSINT stand as a paradox opposite the historical foundations of the Intelligence Community conducting clandestine activities. The statistics of the 21st century explosion of PAI speaks for itself creating the current age of Big Data. Instead of exploiting traditional classified intelligence, tasking the Big Data process, which neither the IC owns nor is clandestine, provides the strongest source to build a strategic warning process.

Big Data intersects with Open Source Intelligence supporting the strategic warning process to power sharper analytical conclusions that will challenge long held assumptions. Traditional intelligence analysis attempts to discover causation within events in the international system. In contrast, Big Data analytics undertakes the search for correlation of events and discovers new insight into challenging strategic questions to open the range of possibilities. An exploration of private industry and non-profit harnessing Big Data demonstrates how this flow of overwhelming information powers this search to answer strategic questions and produces predictive analysis.

How does private industry use Big Data?

A review of the Big Data field reveals the nascent nature of an emerging discipline. Nonetheless, the manner in which industries outside of government utilize Big Data shows the potential this field holds to formulate strategic warning. Cukier and Mayer-Schonberger demonstrate that “at its core, Big Data is about predictions” with its ability to inform industry leaders about trends to anticipate in order to gain a competitive advantage in the marketplace.⁴⁶ The researchers utilize for their strongest arguments the enormous amount of data retail giants Amazon and Walmart generate for analysis. Urging to “let the data speak for itself,” the researchers show through marketplace examples that decision makers must discern the important difference between correlation and causation when analyzing Big Data. In data science, the phrase “correlation does not imply causation” guides statisticians to recognize that the presence of one variable does not mean with certainty it created the other.⁴⁷

Walmart and Amazon analyzing their enormous collection of data provide fascinating examples for Cukier and Mayer-Schonberger to offer their overall argument regarding Big Data. The vast size of Big Data shifts the focus from surveying smaller samples of information that aimed for precision to discovering new strategic insight from large, unstructured data sets.

Shifting focus from small to large with Big Data moves the overall analytical focus away from the search for causality to a “discovery of patterns and correlations in the data that can offer us novel and invaluable insights” not previously seen. The cogent example of flashlights and pop tarts is particularly compelling. In a particular time frame in their Florida stores, Walmart discovered a pattern in their sales data that a spike in the sales of flashlights transpired along with a similar spike in the sales of strawberry pop tarts. In the search for causality, the sales data alone shows one variable may cause the other. Walmart realized that just the data alone each cash register transaction generates is not enough for true Big Data analysis. Along with everything in each customer’s shopping cart, Walmart also added additional data variables including the time of day, geography, demographics, and even the weather. With the additional data, the retailer discovered that a forecasted hurricane preceded the jump in flashlights creating a correlation to an uptick in strawberry pop tarts. In performing correlation analysis instead of searching for causality, Walmart uncovered invaluable insight to properly stock their shelves and position inventory during a hurricane because the storm is the causation that creates the correlation of flashlights and pop tarts. Walmart’s use of Big Data emphasizes the importance of data analysis alongside the statistical compilation of all available information enabling the uncovering of correlations in the data.⁴⁸

In a similar Big Data effort, online retailer Amazon perfected an algorithm to produce recommendations of products that personalizes each individual’s shopping experience tailored to their unique preferences. Almost 15 years in use, the Amazon algorithm known as “item-to-item collaborative filtering” recommends with high accuracy items a unique customer could purchase by scaling against a massive data set. Existing algorithms failed to generate high quality recommendations for Amazon because traditional collaborative filtering attempts to cluster similar customers with their product interests missed correlations. Basing collaborative filtering

on items instead of attempting to match purchases of similar customers “recommends highly correlated similar items, thus recommendation quality is excellent.” Amazon is an online pioneer in generating a tailored shopping experience with the creation of their algorithm that has set the industry standard.⁴⁹

The latest Amazon endeavor clearly demonstrates the power of Big Data to inform predictive analytics by shipping products to customers before their transaction. Amazon patented a system to produce “anticipatory shipping” which will predict items a customer will purchase at a certain time, putting the merchandise in motion towards a location before the sale even occurs. The patented algorithm for anticipatory shipping builds a forecasting model based on prior shopping history, the time a particular item has been with a customer, length of webpage views, links clicked or hovered, and wish lists. Secondary data such as time of year, holidays, large sporting events, geographic location, weather patterns, and demographics are overlaid to enrich the data. Similar to Walmart, correlation analysis predicts the correct items to ship to either a regional fulfillment center or even place on a local truck ready for delivery. Also, anticipatory shipping lessens the impact of the biggest weakness of Amazon when compared to traditional brick and mortar stores, the immediate gratification of a purchase. Optimizing the logistics of supply chain with anticipatory shipping provides a company the size of Amazon a significant opportunity to save money, generate additional revenue, and increase customer satisfaction.⁵⁰

The Big Data lesson from industry seeking to gain a competitive advantage is the strongest predictive technique identifies correlation within events along with recognizing the difference from causation. The search for correlation discovers additional variables in the data not previously uncovered allowing for the assessment of a full range of outcomes. Big Data analytics opens the aperture to discover all possible correlations including items previously missed. Walmart overlaying weather data with retail activity discovered the outcome of

correlated events produced by a storm. Amazon analyzing massive data sets including secondary information instead of only previous customer purchases opened the spectrum of correlated events producing high quality recommendations and enabling anticipatory shipping. Discovering correlated events through Big Data analytics powers a strategic warning process by anticipating all possible outcomes producing high quality assessments for leaders.

How Does the Non-profit Sector use Big Data?

While Amazon and Walmart demonstrate large retailers using correlation analysis on their enormous internally generated datasets, non-profit and non-governmental organizations (NGO) provide a model for collecting publically available information to harness the power of Big Data. Specifically with methods for crowdsourcing and crisis mapping, these organizations tap the power of personal communication technology to penetrate areas of ungoverned spaces experiencing limited statehood to provide alternative forms of governance. In creating alternative governance, several standing assumptions are negated. Wireless communication technology pervades into the most constrained of geopolitical environments, as citizens of failed states experiencing a natural disaster or even large scale violent conflict can fill the state void with innovative communication tools. As wireless technology drastically lessened in cost, cellular devices and smartphones became available to citizens of poor regions, digitally empowering small groups of people to create Big Data of high quality openly available for collection. People managing their lives in areas of limited statehood nullify standing assumptions regarding the perception of projecting power as the economically disadvantaged possess wireless technology enabling governance in these areas while simultaneously creating Big Data.⁵¹

Several examples provide a model for sourcing user-generated information from these areas of limited statehood to inform strategic warning. During the humanitarian disaster created in the wake of the Haiti earthquake in January 2010, the strain placed on a limited government

proved too intense for internal emergency response. Big Data created an avenue for an alternative mode of government as social media content allowed for outsourcing of the emergency response to NGOs and foreign states. Citizens of a failed state experiencing a large-scale humanitarian emergency had the means through personal ownership of wireless technology to project power, supporting the emergence of alternative actors to lead through the emergency.⁵²

Crisis mapping organization Usahidi is a case study of the emergence of an alternative actor providing leadership to create governance. With its origin monitoring Kenyan elections in 2008, Usahidi matured their crisis mapping tradecraft during the Haiti Earthquake of 2010 displaying the power of Big Data. Collecting the information created by Haitians using personal communication technology in the midst of the disaster, Usahidi crowd-sourced social media content posted to platforms such as Twitter, Facebook, and YouTube. Also, the group established a central platform to receive SMS messages directly from cell phones in Haiti. Volunteers with Usahidi created crisis maps from unstructured social media data enabling geospatial visualization of areas with the most damage and people in need empowering decision makers to determine where to direct the appropriate emergency response.⁵³

While allowing people to self-govern within areas of limited statehood, crowdsourcing harnesses the power of Big Data to create strategic intelligence. In the 2011 Libyan uprising, Usahidi answered information requirements for the United Nations with crisis maps tracking an unfolding humanitarian crisis as refugees fleeing violence required an emergency response. The same formula of crowdsourcing social media content created by refugees with wireless technology in their pocket reporting their conditions to social media proved effective in saving lives mitigating dire conditions. The UN levied a crowdsourcing request against Usahidi capabilities who responded with a highly accurate map detailing evacuation routes, movement of refugees with size, locations of conflict, and military movements. The crowdsourcing effort

provided such precise information that Usahidi kept most details behind a firewall so that Libyan forces could not exploit the product for their intelligence purposes.⁵⁴

Building on the strength of the crowdsourcing model, crowdseeding is the next evolution in generating and collecting Big Data. An organization utilizing crowdseeding equips a group of citizens in a specific geographic area of interest with a cellular device and given training on how to send information to an SMS collection platform. The strength of crowdseeding is “the seeds” can be chosen with specificity to ensure a full sample of information is collected across all demographics of the people under consideration. In a crowdseeding pilot, Columbia University partnered with USAID to fund *Voix des Kivus*, a program collecting data from citizens in the province of South Kivu, a war-torn area in the Democratic Republic of the Congo. Translating appropriately to “voice of the Kivus,” three cellphone holders were selected in each individual village; one traditional leader, one woman, and one elected by the village. Each holder received extensive training on how to report specific cases of violence, military incidents, illegal activities, and NGO activities.⁵⁵ An analysis upon completion showed the pilot generated high quality data on the nature of life in the war torn region with low cost and technical barriers. Research lacks for determination of ethical implications for crowdseeding regarding the protection of cellphone holders that ensures all voices within a society are receiving attention.⁵⁶

Social enterprise Spatial Collective is an example of persistent crowdseeding of a target area of concern to generate a constant stream of Big Data. Using commonly available technology, such as mobile devices, GPS units, and digital cameras, personnel from Spatial Collective map the Nairobi slum city of Mathare with the mission to document urgent issues giving them attention, in order to remedy situations where a lack of resources exists. Spatial Collective is facilitating action to supply a solution to urban problems.⁵⁷ When a sewage issue is visualized on their map, a conduit is created for action by either a governmental agency or NGO

to apply solution that will ward off possible disease outbreaks. Crowdfunding creates strategic warning by anticipating issues created by urban poverty before they occur. The process identifies sewage issues before contaminated water creates a cholera outbreak and adjudicates property disputes before outbreaks of violence occur.⁵⁸

NGO utilization of Big Data presents a model for emulation especially by the Intelligence Community. Unlike large volume retailers, the federal government will not generate data of intelligence value that will inform strategic warning. NGOs operate in a constrained resource environment with small staffs and limited budgets. Using methods such as crowdsourcing or crowdfunding develop a readily available persistent sensor for Big Data collection on targeted areas of interest by an organization with limited resources. NGOs penetrating constrained environments with areas of limited statehood using innovative technologies directly support key components of strategic warning. A feature of the strategic warning process is the analyst engaging all possible intelligence in an iterative fashion over an extended amount of time. An innovative approach such as crowdfunding collects intelligence from an austere environment providing Big Data source from an unlikely location. Requirements for strategic warning from numerous locations for the IC include locations that innovative technologies are reaching for the first time. Social media data meets all “four V’s” of the Big Data: volume, velocity, variety, and veracity. Social media content promises to pour from emerging markets in Asia and Africa where intelligence collection currently remains limited.

Proposed Courses of Actions to Build a Strategic Warning Process

This paper has established the definition of strategic warning, tracing how it has been used historically, arriving to its present day state with emphasis on intelligence estimates resting in the NIC and with DIA for the Department of Defense. With the era of Big Data, industry and non-profit organizations demonstrate a model on how the deluge of PAI can be harnessed for

predictive analytics. Several courses of action are proposed to build a strategic warning process in the IC using Big Data analytics compiled from publically available information.

Build on DARPA's Big Data Program

Federal Government agencies and the Intelligence Community are not ignoring Big Data. In 2012, in an effort to tackle the dynamic of Big Data, the Department of Defense turned to the Defense Advanced Research Projects Agency (DARPA) to develop solutions. DARPA is an agency of the Department of Defense charged with researching and developing innovative projects in an attempt to push the limits of science and technology in an effort to meet military requirements. DARPA has been responsible for developing several significant technological achievements, such as GPS, the Internet, and stealth capabilities. In an effort to meet the challenge of Big Data, DARPA established a program known as XDATA. As current programs in DOD struggle to scale effectively the enormous volume of data it currently collects and stores then in turn provide data analysis, XDATA hopes to answer these big data challenges with techniques and tools for processing and analyzing large, unstructured data sets. XDATA is building an open source software catalog in hopes of discovering solutions to scale effectively big data volume and characteristics. The XDATA team consists of computer developers that write complex code to build tools and data scientists that are developing advanced algorithms all built on an open source software library.⁵⁹

The DARPA open catalog compiles all of its Big Data work with an organized list of select software and peer-reviewed publications produced by partners from academia and industry contributing to the XDATA program. The intent of DARPA is to build a community around solving Big Data problems supported with DOD funding that will produce software, publications, and experimental approaches to data using the scientific method. DARPA's

strategy is complete transparency with all results, software, and coding in order to increase the impact of the program across the partner sectors of industry, academia, and government.⁶⁰

A review of DARPA's Big Data open catalog of how the program is managed shows this effort has data science in mind, not traditional intelligence analysis. The research of DARPA is managed by six technical offices. Their explorations into Big Data are spearheaded by their Information Innovation Office (I2O), whose mission is to develop state-of-the-art programs for information science and technology with a focus on three main areas: cyber, analytics, and symbiosis. DARPA placing Big Data in their I2O branch shows it believes Big Data is an issue for information technology to solve.⁶¹

Similar to Usahidi and Spatial Collective, DARPA's Big Data program attempts crowdsourcing and crowdseeding, but the efforts are treated through the lens of data science rather than answering intelligence requirements. Crowd-Sourced Formal Verification (CSFV) is a program that is investigating if large numbers of common individuals with no expert knowledge can perform formal programming verification more timely and effectively than conventional processes with experts powered by machine learning. Researchers at the University of Washington participating in the program developed a method attempting to make verification more effective by reducing the skills needed for coding verification along with increasing the pool of participants. To achieve simplification and wide participation, the researchers designed an engaging visual puzzle game intuitive enough that an average person could gain expertise by solving it. The goal is to verify the correct programming of game software by crowdsourcing students and workers during their free time while waiting at mass transit stations. Instead of relying on a small pool of expert software engineers, crowdsourcing produced a wide audience of people who enjoy the challenge of a puzzle and could simultaneously verify the programming code of the game.

Conducted by researchers from the Computer Science and Engineering department, the complex study analyzes the results of crowdsourcing by examining the discovery of mistakes and the mapping of security vulnerabilities. One conclusion is enlightening: “We believe that humans may have an edge over automated systems in certain situations, notably when the program is not verifiable. A program is unverifiable when it has a bug, or when it is correct for reasons that are beyond the power of the verification system.”⁶² When an automated system cannot detect a bug or verify it is correct, human interaction can provide verification, even from a non-expert. Yet, no explanation or additional exploration is offered by the UW study into this finding. Data science does not offer an answer to this dynamic where analysis through a socio-cultural lens would be helpful to explain why crowdsourcing non-experts are fruitful.

With applications for crowdseeding, GRAPHS is a Big Data program at DARPA with the mission of developing models that predict anomalies and exceptional events in very large constantly evolving networks. To achieve this mission, the program develops statistical methods powered by algorithms to detect these network events with a focus on the quantitative and structural components of large networks. DARPA is actively funding software development and peer-reviewed research by partners in academia and industry. A review of the primary data from GRAPHS reveals where DOD is headed with predictive analysis of large networks.

One of the dynamics the program explores is if the individuals in a large dynamic network best situated to diffuse information can be identified by inquiring only a few members of the network. In this study, a group of economists uses a theory known “diffusion centrality” on data collected from several communities in India to determine if members of the network accurately determine which people are the best diffusers of information across the community and how far the information diffuses over a certain period of time. The diffusion of information is tracked across the village using a small number of randomized seeds who are tasked to spread

the information. If villagers select a strong diffuser of information identified as a “good gossip,” information spread by an increase of 65%. In addition, a polling of all adults in the 213 communities under consideration determined the best gossip seeds to spread information. The researchers then split the communities into thirds, seeding information with randomized members in 71 networks, the leaders and elders in 71 networks, and finally utilized the best gossip seeds as nominated in the remaining 71 networks. Research showed that information traveled the least in the randomized set, further with leaders and elders, yet the identified best diffusers of information doubled the rate of leaders and elders. An interesting trend emerges in identified gossips versus community leaders – only 8% of the nominated best gossips were also leaders or elders, meaning that 92% of the best diffusers of information are not identified as a network leader or elder of the village.⁶³

The solution proposed is for the Intelligence Community to build upon what DARPA’s Big Data program is establishing. The open program model in the XDATA program DARPA utilizes with developing advanced algorithms in an open source code library is perfect for Big Data analytics. The world of Big Data evolves rapidly at such a pace, the Intelligence Community will struggle to maintain pace. Establishing an open source code library will improve Big Data analytics in practical real-time. Ultimately, open source code crowdsources Big Data analytics while encouraging participation from academia and NGOs as these institutions have a stake in the program’s success. A large open source coding community allows for the discovery of bugs and applying a remedy early in the process. Although, the open source model DARPA utilizes is anathema to the Intelligence Community and its traditional view of classification levels. A cultural shift of the IC will have to continue to take place embracing the success of industry and NGOs of building strategic warning from unclassified Big Data.

DARPA emphasizes data science as a critical component of the Big Data equation. Yet, these fields with their tradecraft are only part of the equation. An algorithm programmed by a data scientist is left without an explanation in the crowdseeding study on why the senior leaders of an Indian village diffuse information so poorly through their community. To provide a full assessment, an analyst with Indian expertise is needed that can apply socio-cultural analysis. As Grabo argues, analytical rigor is crucial for providing strategic warning by going deep into a particular topic over an extended time. In order to support the key components of the strategic warning process, the data scientist and the intelligence analyst must work in unison each applying their tradecraft. The DARPA establishes the groundwork for data science, providing the opportunity for intelligence analysis to complete the strategic warning process.

Where in the Intelligence Community should a Big Data strategic warning process reside?

As outlined in this paper, tracing the historical evolution from the founding of the IC to present shows a primary mission for strategic warning residing with DIA and the NIC. These are two different intelligence organizations as the NIC is a focused analytical effort with senior policymakers as their primary customer, while DIA focuses on coordinating intelligence functions for the Defense Intelligence Enterprise with the Pentagon as their primary customer. However, the ultimate answer of where a strategic warning process powered by Big Data should reside is throughout the IC, not just with one agency. Since strategic warning historically sits with DIA and the NIC, these agencies would be among XDATA's primary customers as the National Security Council, the White House, and the Pentagon regularly lean into the NIC and DIA for intelligence estimates.

A strategic warning process diffuse throughout the IC sharing Big Data analytics would allow each intelligence agency to build strategic warning products tailored to their customers. As DARPA's data scientist develop algorithms producing Big Data analytics, their results can be

spread across the IC for engagement. Using a community model also provides the ability to build a tradecraft group with representation from all 16 intelligence agencies to provide feedback to the team of data scientists and sharpen analytical models. A community driven approach also allows for shared cost and data storage. A survey of industry solutions for Big Data collection, storage, and retrieval shows that requirements are readily available in commercial off the shelf product that will require budgets resources to purchase. The industry standards for Big Data collection and storage are at the forefront of innovation and address all requirements for volume, scope, and natural language processing.⁶⁴

The Pentagon is making heavy investments in the resourcing and development of Big Data for storage and retrieval from these commercial products. The FY 2017 procurement budget request for Big Data stands close to \$600 million.⁶⁵ This procurement request purchases items such as Dataminr, a tool that converts a persistent stream of unstructured social media data into actionable alerts tailored to each user's individual requirements.⁶⁶ A community model would not only allow for cost sharing, but for the sharing of tradecraft that powers these Big Data tools across the IC.

A data scientist with their algorithm will not be sufficient to produce strategic warning. Walmart needed an analyst to discern correlation and causality in examining flashlights, pop tarts, and hurricanes. Analysis of the crowdsourced data Ushahidi utilized in production of crisis maps is required to determine veracity. Crowdfunding in Kenya necessitates human critical thinking to adjudicate issues when collecting data. For Big Data to produce strategic warning, the data scientist and the intelligence analyst must work in tandem.

Create a Big Data analytical framework for a Strategic Warning process

Strategic warning is a function of analysis and needs an analytical framework for analysts to apply in order for Big Data to power predictive intelligence. Several items are needed to build

this framework. Big Data and strategic warning lack a doctrinal definition. For Big Data, information technology is conflated with the analytical field when attempting to reach a definition. The phrase is sometimes used to describe a technical architecture to collect, store, and retrieve large volumes of data. In some instances, the phrase is used to describe the work of data scientists developing algorithms.⁶⁷ A doctrinal definition needs to describe all aspects of Big Data. The RAND definition meets this need as it borrows the industry standard of the “four V’s of Big Data” with volume, variety, velocity, and veracity.⁶⁸ The “four V’s” definition addresses technology with the overwhelming volume and diversity along with the mechanics needed to mine the data. Veracity addresses intelligence analysis concerns as determining reliability and credibility of the data pushes the bounds of traditional analytics.

As to a strategic warning definition, Grabo readily admits that a single definition strategic warning is not established. Conceptually, the analytic community is not confused on what constitutes strategic warning as the definition is apparent. Although, placing into doctrine a definition for strategic warning would help establish an analytical framework by clearly articulating what constitutes strategic warning. Joint Publication 2-0 *Joint Intelligence* provides a strategic warning definition already placed in doctrine which communicates warning provides a distinct communication to a leader about a threat allows the leader time and options to counter and mitigate the threat. This definition can serve as a template to build a definition for the national IC, especially as it distinguishes the two types of warning, emerging and enduring.⁶⁹

After establishing the definition, how to achieve strategic warning needs articulation with parameters and goals. Success in strategic analysis is not judged by producing precise predictions that later come true. A strategic warning process analyzes a range of possible outcomes with provision of courses of action that a decision maker could pursue. Strategic warning is an iterative process transpiring over an extended period of time. As Grabo encourages, the strategic

analyst going deep into the collected intelligence over many years is imperative for the warning process. Rigorous analysis negating standing assumptions with flexible frameworks will provide recognition of true signals out of the noise as Wohlstetter demonstrates is essential. The intense noise and enormous volume of Big Data necessitates iterative analysis recognizing truth within boundless information. Veracity is the fourth V in the definition of Big Data, which means literally truthfulness. Producing the fourth V will be the hard work and analytical rigor of the intelligence professional.

Conclusion and Recommendations for Future Research

IBM CEO Virginia Rometty in a speech to the Council on Foreign Relations argued, “Big Data is the new natural resource” and will “force every company, country, government entity, to actually become an authentic organization.”⁷⁰ Just as naturally occurring minerals and forests can be harvested for gain, Big Data occurs naturally and is available for harvesting to produce an economic or intellectual gain. This natural resource needs to properly protected, harvested, and utilized by government agencies, especially by the Intelligence Community.

Considering how to manage properly with intent to harvest the new natural resource of Big Data, proposed are questions for future research. The clandestine activity of creating and tasking sources for collecting intelligence is familiar tradecraft for the Intelligence Community. In a similar fashion, can intelligence activities proactively crowdseed areas of interest, similar to developing HUMINT sources over an extended timeframe? For example, a known terror group operates in an area of limited statehood that experiences low cell phone penetration. A crowdseeding effort potentially could allow innocent non-combatants the ability to self-govern through collective action utilizing personal communication technology, similar to Spatial Collective’s effort in Nairobi. The content generated is then crowdsourced to collect intelligence on terrorist activities in an area previously without penetration. An area open for future research

is the possibility of the IC establishing a crowdsourcing process for intelligence collection. New research on crowdsourcing should also include exploration of the possible ethical dilemmas surrounding exploitation of people under duress for intelligence collection.

This study argues that analytical tradecraft is a vital piece along with data science to produce a strategic warning process using Big Data. Can the Intelligence Community produce a body of work on the analytical tradecraft of Big Data producing strategic warning? Several institutions in the IC at the community level maintain analytical standards with critical skills for analysis. As the use of Big Data grows, training organizations such as the Sherman Kent for Intelligence Analysis at the CIA and the Joint Military Intelligence Training Center could build curriculum that maintain training standards for strategic analysis using Big Data as the intelligence source. New research on training for strategic warning could also explore whether it is achievable to produce a cadre of intelligence analysts focused solely on strategic warning. Grabo argues analysts working in a specific topic over an extended time produce effective strategic analysis. Some intelligence agencies may have difficulties devoting a team of analysts only to strategic warning given resource constraints.

A question to consider when collecting Big Data involves the possible dilemmas for issues of intelligence on U.S. persons. Avoiding collection of information in open sources generated by U.S. persons may be close to impossible. The laws regulating traditional intelligence collection on U.S. citizens also apply to collection from open sources, such as Executive Order 12333 and Department of Defense Regulation 5240.1-R. In the forming age of Big Data, legal scholars and government leaders may possibly need to contribute additional guidance and legal regulation for collection activities. Government policies must emphasize solutions to constitutional concerns of right to privacy raised by Big Data collection.

¹ Svetlana Sicular, “Gartner’s Big Data Definition Consists of Three Parts, Not to be Confused with Three ‘V’s,” *Forbes*, March 27, 2013,

<https://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/>

² Stew Magnuson, “Military ‘Swimming in Sensors and Drowning in Data,’” *National Defense*, January 2010, <http://www.nationaldefensemagazine.org/archive/2010/January/Pages/default.aspx>

³ M.G. Siegler, “Eric Schmidt: Every 2 Days We Create As Much Information As We Did Up To 2003,” *Tech Crunch*, August 4, 2010. <https://techcrunch.com/2010/08/04/schmidt-data/>

⁴ *National Defense Authorization Act for Fiscal Year 2006*, Public Law 109-163, HR 1815, 109th Congress, January 6, 2006, Section 931(a), page 119.

⁵ Isaac Porche, III, Bradley Wilson, Erin-Elizabeth Johnson, Shane Tierney, and Evan Saltzman, “Data Flood: Helping the Navy Address the Rising Tide of Sensor Information,” *The RAND Corporation* (2014): http://www.rand.org/pubs/research_reports/RR315.html

⁶ Mary McCarthy, “The National Warning System: Striving for an Elusive Goal,” *Defense Intelligence Journal* 3 (1994): 5-19.

⁷ The Office of the Director of National Intelligence, *The National Intelligence Strategy of the United States of America* (Washington, DC, 2014), https://www.dni.gov/files/documents/2014_NIS_Publication.pdf.

⁸ Defense Intelligence Agency, *Defense Intelligence Agency Strategy* (Washington, DC, 2016), http://www.dia.mil/Portals/27/Documents/About/2016_DIA_Strategy.pdf.

⁹ The U.S. Department of Justice, *Office of the Attorney General, United States Department of Justice Fiscal Years 2014-2018 Strategic Plan* (Washington, DC, 2014), <https://www.justice.gov/sites/default/files/jmd/legacy/2014/02/28/doj-fy-2014-2018-strategic-plan.pdf>

¹⁰ The White House, *The National Security Strategy of the United States of America* (Washington, DC, February 2015), <http://nssarchive.us/wp-content/uploads/2015/02/2015.pdf>.

¹¹ Cynthia M. Grabo, “Anticipating Surprise: Analysis for Strategic Warning,” *Center for Strategic Intelligence Research* (December 2002): http://www.ni-u.edu/ni_press/pdf/Anticipating_Surprise_Analysis.pdf

¹² McCarthy, “The National Warning System: Striving for an Elusive Goal,” 5-19.

¹³ Jeffrey R. Cooper, “Curing Analytic Pathologies: Pathways to Improved Intelligence Analysis,” *Center for the Study of Intelligence, Central Intelligence Agency* (December 2005): https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/curing-analytic-pathologies-pathways-to-improved-intelligence-analysis-1/analytic_pathologies_report.pdf

¹⁴ Grabo, “Anticipating Surprise: Analysis for Strategic Warning,” 4-6.

¹⁵ Michael Warner, ed., “Central Intelligence: Origin and Evolution,” *Center for the Study of Intelligence, Central Intelligence Agency*, (2001): https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Origin_and_Evolution.pdf.

¹⁶ National Security Council Intelligence Directive No. 1, Washington DC, January 19, 1950. Foreign Relations of the United States, 1945-1950, Emergence of the Intelligence Establishment, State Department Office of the Historian. <https://history.state.gov/historicaldocuments/frus1945-50Intel/d432>

¹⁷ National Security Council Intelligence Directive No. 5, Washington DC, December 12, 1947. Foreign Relations of the United States, 1945-1950, Emergence of the Intelligence Establishment,

State Department Office of the Historian. <https://history.state.gov/historicaldocuments/frus1945-50Intel/d423>

¹⁸ Jeffrey Richelson, *A Century of Spies: Intelligence in the Twentieth Century* (New York: Oxford University Press, 1994), p. 293-297.

¹⁹ P.K. Rose. "Two Strategic Intelligence Mistakes in Korea, 1950," *Studies in Intelligence, Center for the Study of Intelligence, Central Intelligence Agency*, (Fall-Winter 2001): https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/fall_winter_2001/article06.html.

²⁰ Ibid.

²¹ Michael Warner and J. Kenneth McDonald, *US Intelligence Community Reform Studies Since 1947* (Washington, DC: Center for the Study of Intelligence, April 2005), 11-13. <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/US%20Intelligence%20Community%20Reform%20Studies%20Since%201947.pdf>

²² National Security Council Intelligence Directive 5511, Washington DC, February 14, 1955. Foreign Relations of the United States, 1950-1955, The Intelligence Community, 1950-1955, State Department Office of the Historian. <https://history.state.gov/historicaldocuments/frus1950-55Intel/d207>

²³ James H. Doolittle et al., "Report on the Covert Activities of the Central Intelligence Agency," *The Central Intelligence Agency, CIA Library*, (July 26, 1954): <https://www.cia.gov/library/readingroom/docs/CIA-RDP86B00269R000100040001-5.pdf>

²⁴ Warner and McDonald, *US Intelligence Community Reform Studies Since 1947*, 15.

²⁵ Christopher M. Andrew, "American Presidents and Their Intelligence Communities," in *Intelligence and the National Security Strategist*, ed. Roger Z. George and Robert D. Kline. (New York: Rowman & Littlefield Publishers, Inc., 2006), 431- 445.

²⁶ Roberta Wohlstetter, "Cuba and Pearl Harbor: Hindsight and Foresight," The RAND Corporation (April 1965): https://www.rand.org/content/dam/rand/pubs/research_memoranda/2007/RM4328.pdf

²⁷ Ibid., 10-12.

²⁸ Defense Intelligence Agency, *A History of the Defense Intelligence Agency* (Washington, DC: DIA Office of Historical Research, 2007), https://fas.org/irp/dia/dia_history_2007.pdf.

²⁹ *Worldwide Threat Assessment : Testimony before the Armed Services Committee*, 113th Cong. (February 3, 2015) (statement of Lieutenant General Vincent R. Stewart, Director, Defense Intelligence Agency). <http://www.dia.mil/News/Speeches-and-Testimonies/Article-View/Article/567087/worldwide-threat-assessment/>

³⁰ Cheryl Pellerin, "DIA Five Year Plan Updates Strategic Warning Mission," American Forces Press Service, July 18, 2012 <https://www.dni.gov/index.php/newsroom/ic-in-the-news/187-ic-in-the-news-2012/583-dia-five-year-plan-updates-strategic-warning-mission>.

³¹ *The National Intelligence Council, Title 50 U.S.Code §403-3b(b)(1)*.

³² Richard A. Best, Jr., *The National Intelligence Council: Issues and Options for Congress*, CRS Report for Congress R40505 (Washington, DC: Congressional Research Service, January 10, 2011), 5-6.

https://digital.library.unt.edu/ark:/67531/metadc31463/m1/1/high_res_d/R40505_2011Jan10.pdf

³³ Ibid, 7.

³⁴ Anthony Olcott, *Open Source Intelligence in a Networked World* (New York: Bloomsbury Academic, 2012), 25-27.

³⁵ IBM, "Bringing Big Data to the Enterprise" <https://www-01.ibm.com/software/in/data/bigdata/> (accessed February 1, 2017).

-
- ³⁶ Pew Research Center, “Mobile Fact Sheets: Internet/Broadband | Social Media,” January 12, 2017, <http://www.pewinternet.org/fact-sheet/mobile/>.
- ³⁷ Stephanie Mlot, “70 Percent of Population Will Have Smartphones by 2020,” *PC Magazine*, June 3, 2015, <http://www.pcmag.com/article2/0,2817,2485277,00.asp>
- ³⁸ National Intelligence Council, Office of the Director of National Intelligence, *Global Trends: Paradox of Progress*. (Washington, DC: January 9, 2017), <https://www.dni.gov/index.php/about/organization/national-intelligence-council-global-trends>.
- ³⁹ Osterman Research, Inc. “Key E-Discovery Issues in 2011,” April 2011, http://www.ostermanresearch.com/whitepapers/or_or0411a.pdf.
- ⁴⁰ Sharath Srinivasan, “FrontlineSMS, Mobile-for-Development, and the ‘Long Tail’ of Governance,” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Steven Livingston and Gregor Walter-Drop (Oxford: Oxford University Press, 2014), 79-97.
- ⁴¹ W. Lance Bennett, “The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation,” *The Annals of the American Academy of Political and Social Science* Vol 644, Issue 1 (2012): 20-39.
- ⁴² The Intelligence Reform and Terrorism Prevention Act of 2004 (IRTPA) Public Law 108-458 (December 17, 2004). <https://it.ojp.gov/PrivacyLiberty/authorities/statutes/1282>
- ⁴³ *National Defense Authorization Act for Fiscal Year 2006*, Public Law 109-163, HR 1815, 109th Congress, January 6, 2006, Section 931(a), page 119.
- ⁴⁴ United States Department of Defense, *Joint Intelligence*, Joint Publication 2-0 (Washington, DC: Government Printing Office, October 22, 2013). http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf.
- ⁴⁵ Porche, Wilson, Johnson, Tierney, and Saltzman. *Data Flood: Helping the Navy Address the Rising Tide of Sensor Information*, 2-3.
- ⁴⁶ Viktor Mayer-Schonberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. (New York: Houghton Mifflin Harcourt, 2013).
- ⁴⁷ Ibid, 50-55.
- ⁴⁸ Ibid, 53-54.
- ⁴⁹ Greg Linden, Brent Smith, and Jeremy York, “Amazon.com Recommendations: Item-to-Item Collaborative Filtering,” *IEEE Internet Computing* (January-February 2003): 76-80. <https://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf>
- ⁵⁰ Lydia DePillis, “Amazon Wants to Send Stuff before You Order It. Are other retailers doomed?,” *The Washington Post*, January 30, 2014, <https://www.washingtonpost.com/news/wonk/wp/2014/01/30/amazon-wants-to-send-stuff-before-you-order-it-are-other-retailers-doomed/>
- ⁵¹ Steven Livingston and Gregor Walter-Drop, “Introduction,” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Steven Livingston and Gregor Walter-Drop (Oxford: Oxford University Press, 2014), 1-16.
- ⁵² Gregory Asmolov, “Natural Disasters and Alternative Modes of Governance: The Role of Social Networks and Crowdsourcing Platforms in Russia,” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Steven Livingston and Gregor Walter-Drop (Oxford: Oxford University Press, 2014), 98-114.
- ⁵³ Penn State Public Broadcasting, “The Geospatial Revolution,” 2010, documentary, <http://geospatialrevolution.psu.edu/> <https://www.youtube.com/watch?v=gxCEb5Cv4N>
- ⁵⁴ Patrick Meier, “Crisis Mapping in Areas of Limited Statehood,” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Steven Livingston and Gregor Walter-Drop (Oxford: Oxford University Press, 2014), 130-143.

-
- ⁵⁵ Peter van der Windt, "From Crowdsourcing to Crowdfunding: The Cutting Edge of Empowerment?," in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Steven Livingston and Gregor Walter-Drop (Oxford: Oxford University Press, 2014), 144-155.
- ⁵⁶ Peter van der Windt and Macartan Humphreys, "Crowdfunding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time," *Journal of Conflict Resolution* Vol 60, Issue 4 (2016): 24-16. <http://journals.sagepub.com/doi/pdf/10.1177/0022002714553104>
- ⁵⁷ Primoz Kovacic and Jamie Lundine, "Mapping Kibera. Empowering Slum Residents by ICT," in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Steven Livingston and Gregor Walter-Drop (Oxford: Oxford University Press, 2014), 115-128.
- ⁵⁸ Gregory Warner, "In Kenya, Using Tech to Put An 'Invisible' Slum on the Map," NPR, Washington DC: July 17, 2013. <http://www.npr.org/sections/parallels/2013/07/17/202656235/in-kenya-using-tech-to-put-an-invisible-slum-on-the-map>
- ⁵⁹ Defense Advanced Research Projects Agency, "The XDATA Program," [darpa.mil, http://www.darpa.mil/program/xdata](http://www.darpa.mil/program/xdata), (accessed January 4, 2017)
- ⁶⁰ Defense Advanced Research Projects Agency, "The DARPA Open Catalog," [darpa.mil, http://opencatalog.darpa.mil/](http://opencatalog.darpa.mil/), (accessed January 4, 2017)
- ⁶¹ Defense Advanced Research Projects Agency, "DARPA Offices," [darpa.mil, http://www.darpa.mil/about-us/offices](http://www.darpa.mil/about-us/offices), (accessed January 4, 2017)
- ⁶² Ibid.
- ⁶³ Abhijit V. Banerjee, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson, "Gossip: Identifying Central Individuals in a Social Network," (February 14, 2016): MIT Department of Economics Working Paper No. 14-15: <https://ssrn.com/abstract=2425379> or <http://dx.doi.org/10.2139/ssrn.2425379>
- ⁶⁴ Joe McKendrick, "First, There Was Big Data; Now Comes Big Storage." Database Trends and Applications 26.2 (June 2012): 6-8, 10. ImmixGroup, "FY16 Government Panel Discussion on Big Data Opportunities at Federal Agencies," Panel Discussion, Washington, DC, June 18, 2015. <https://www.immixgroup.com/Resources/Videos/FY16-Government-Panel-Discussion-on-Big-Data-Opportunities-at-Federal-Agencies/>
- ⁶⁵ The U.S. Department of Defense, *Fiscal Year 2017 Defense Budget Request* (Washington, DC: Department of Defense, 2016), <http://comptroller.defense.gov/Budget-Materials/Budget2017/>.
- ⁶⁶ U.S. Department of Defense, *Intent to Sole Source to Dataminr, Inc for Data Analytics Software License*. (Washington, DC: Department of Defense Acquisition Directorate, 2016), Solicitation Number: KFL0812160249 <https://www.fbo.gov/index?s=opportunity&mode=form&id=f81fd391d15c725e82e84f574bdca8e7>
- ⁶⁷ Porche, Isaac R., I,II, Bradley Wilson, Erin-Elizabeth Johnson, Shane Tierney, and Evan Saltzman. *Data Flood: Helping the Navy Address the Rising Tide of Sensor Information*, 2-3.
- ⁶⁸ "The Four V's of Big Data," International Business Machines, Inc. Infographic, <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>. Retrieved January 8, 2017.
- ⁶⁹ *Joint Intelligence*, I-18.
- ⁷⁰ Ginni Rometty, "A Conversation with Ginni Rometty," Discussion with Q&A, Council on Foreign Relations 2013 Corporate Conference, New York City, March 7, 2013. <http://www.cfr.org/world/conversation-ginni-rometty/p35497>

Bibliography

- Andrew, Christopher M. "American Presidents and Their Intelligence Communities." in *Intelligence and the National Security Strategist*. ed. Roger Z. George and Robert D. Kline. New York: Rowman & Littlefield Publishers, Inc., 2006.
- Asmolov, Gregory. "Natural Disasters and Alternative Modes of Governance: The Role of Social Networks and Crowdsourcing Platforms in Russia." in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. ed. Steven Livingston and Gregor Walter-Drop. Oxford: Oxford University Press, 2014.
- Banerjee, Abhijit V. and Chandrasekhar, Arun G., Duflo, Esther and Jackson, Matthew O. "Gossip: Identifying Central Individuals in a Social Network." (February 14, 2016). MIT Department of Economics Working Paper No. 14-15. <https://ssrn.com/abstract=2425379> or <http://dx.doi.org/10.2139/ssrn.2425379>
- Bennett, W. Lance. "The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation." *The Annals of the American Academy of Political and Social Science*. Vol 644. Issue 1. 2012.
- Best, Richard A., Jr. *The National Intelligence Council: Issues and Options for Congress*. CRS Report for Congress R40505. Washington, DC: Congressional Research Service, January 10, 2011. https://digital.library.unt.edu/ark:/67531/metadc31463/m1/1/high_res_d/R40505_2011Jan10.pdf
- Cooper, Jeffrey R. "Curing Analytic Pathologies: Pathways to Improved Intelligence Analysis." *Center for the Study of Intelligence, Central Intelligence Agency*. December 2005. https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/curing-analytic-pathologies-pathways-to-improved-intelligence-analysis-1/analytic_pathologies_report.pdf
- Defense Advanced Research Projects Agency. "The XDATA Program." [darpa.mil](http://www.darpa.mil/program/xdata). <http://www.darpa.mil/program/xdata>. (accessed January 4, 2017).
- Defense Advanced Research Projects Agency. "The DARPA Open Catalog." [darpa.mil](http://opencatalog.darpa.mil) <http://opencatalog.darpa.mil>. (accessed January 4, 2017).
- Defense Advanced Research Projects Agency, "DARPA Offices," [darpa.mil](http://www.darpa.mil/about-us/offices), <http://www.darpa.mil/about-us/offices>, (accessed January 4, 2017)
- Defense Intelligence Agency. *Defense Intelligence Agency Strategy*. Washington, DC, 2016. http://www.dia.mil/Portals/27/Documents/About/2016_DIA_Strategy.pdf.
- Defense Intelligence Agency. *A History of the Defense Intelligence Agency*. Washington, DC, 2007. https://fas.org/irp/dia/dia_history_2007.pdf.
- DePillis, Lydia. "Amazon Wants to Send Stuff before You Order It. Are other retailers

-
- doomed?” The Washington Post. January 30, 2014.
<https://www.washingtonpost.com/news/wonk/wp/2014/01/30/amazon-wants-to-send-stuff-before-you-order-it-are-other-retailers-doomed/>.
- Dietl, Werner, Stephanie Dietzel, Michael D. Ernst, Nathaniel Mote, Brian Walker, Seth Cooper, Timothy Pavlik, and Zoran Popovic. “Verification Games: Making Verification Fun.” The DARPA XDATA Program. June 12, 2012.
<http://homes.cs.washington.edu/~mernst/pubs/verigames-fffjp2012.pdf>
- Doolittle, James H., William B. Franke, Morris Hadley, and William D. Pawley. “Report on the Covert Activities of the Central Intelligence Agency.” *The Central Intelligence Agency, CIA Library*. July 26, 1954. <https://www.cia.gov/library/readingroom/docs/CIA-RDP86B00269R000100040001-5.pdf>
- Grabo, Cynthia M. “Anticipating Surprise: Analysis for Strategic Warning.” *Joint Military Intelligence College, Center for Strategic Intelligence Research*. December 2002.
http://www.ni-u.edu/ni_press/pdf/Anticipating_Surprise_Analysis.pdf
- IBM, “Bringing Big Data to the Enterprise” <https://www-01.ibm.com/software/in/data/bigdata/>
- The Intelligence Reform and Terrorism Prevention Act of 2004. Public Law 108-458. December 17, 2004. <https://it.ojp.gov/PrivacyLiberty/authorities/statutes/1282>
- Kovacic, Primoz and Jamie Lundine. “Mapping Kiberia. Empowering Slum Residents by ICT.” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. ed. Steven Livingston and Gregor Walter-Drop. Oxford: Oxford University Press, 2014.
- Linden, Greg, Brent Smith, and Jeremy York. “Amazon.com Recommendations: Item-to-Item Collaborative Filtering.” *IEEE Internet Computing*. January-February 2003.
<https://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf>
- Livingston, Steven and Gregor Walter-Drop. “Introduction.” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. ed. Steven Livingston and Gregor Walter-Drop. Oxford: Oxford University Press, 2014.
- McCarthy, Mary. “The National Warning System: Striving for an Elusive Goal,” *Defense Intelligence Journal* 3 (1994): 5-19.
- Magnuson, Stew Magnuson. “Military ‘Swimming in Sensors and Drowning in Data.’” *National Defense*, January 2010.
<http://www.nationaldefensemagazine.org/archive/2010/January/Pages/default.aspx>
- Mayer-Schonberger, Viktor and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin Harcourt, 2013.
- Meier, Patrick Meier. “Crisis Mapping in Areas of Limited Statehood.” in *Bits and Atoms:*

-
- Information and Communication Technology in Areas of Limited Statehood*. ed. Steven Livingston and Gregor Walter-Drop. Oxford: Oxford University Press, 2014.
- Mlot, Stephanie. "70 Percent of Population Will Have Smartphones by 2020." *PC Magazine*. June 3, 2015. <http://www.pcmag.com/article2/0,2817,2485277,00.asp>
- National Defense Authorization Act for Fiscal Year 2006*, Public Law 109-163, HR 1815, 109th Congress, January 6, 2006, Section 931(a), page 119.
- National Intelligence Council. *Global Trends: Paradox of Progress*. Washington, DC: January 9, 2017. <https://www.dni.gov/index.php/about/organization/national-intelligence-council-global-trends>.
- Olcott, Anthony. *Open Source Intelligence in a Networked World*. New York: Bloomsbury Academic, 2012.
- The Office of the Director of National Intelligence. *The National Intelligence Strategy of the United States of America*. Washington, DC, 2014. https://www.dni.gov/files/documents/2014_NIS_Publication.pdf.
- The Office of the Historian, State Department. Foreign Relations of the United States, 1945-1950, *Emergence of the Intelligence Establishment*, Washington, DC. <https://history.state.gov/historicaldocuments/frus1945-50Intel/preface>
- Osterman Research, Inc. "Key E-Discovery Issues in 2011." April 2011. http://www.ostermanresearch.com/whitepapers/or_or0411a.pdf.
- Pellerin, Cheryl. "DIA Five Year Plan Updates Strategic Warning Mission." American Forces Press Service. July 18, 2012. <https://www.dni.gov/index.php/newsroom/ic-in-the-news/187-ic-in-the-news-2012/583-dia-five-year-plan-updates-strategic-warning-mission>.
- Penn State Public Broadcasting, "The Geospatial Revolution," 2010, documentary, <http://geospatialrevolution.psu.edu/> <https://www.youtube.com/watch?v=gxCEb5Cv4Nk>
- Pew Research Center. "Mobile Fact Sheets: Internet/Broadband | Social Media." January 12, 2017. <http://www.pewinternet.org/fact-sheet/mobile/>.
- Porche, Isaac III, Bradley Wilson, Erin-Elizabeth Johnson, Shane Tierney, and Evan Saltzman. "Data Flood: Helping the Navy Address the Rising Tide of Sensor Information," *The RAND Corporation* (2014): http://www.rand.org/pubs/research_reports/RR315.html
- Richelson, Jeffrey. *A Century of Spies: Intelligence in the Twentieth Century*. New York: Oxford University Press, 1994.

Rometty, Ginni “A Conversation with Ginni Rometty.” Discussion with Q&A. Council on Foreign Relations 2013 Corporate Conference. New York City. March 7, 2013.
<http://www.cfr.org/world/conversation-ginni-rometty/p35497>.

Rose, P.K. “Two Strategic Intelligence Mistakes in Korea, 1950.” *Center for the Study of Intelligence, Central Intelligence Agency*. Fall-Winter 2001.
https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/fall_winter_2001/article06.html.

Srinivasan, Sharath. “FrontlineSMS, Mobile-for-Development, and the ‘Long Tail’ of Governance.” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. ed. Steven Livingston and Gregor Walter-Drop. Oxford: Oxford University Press, 2014.

Sicular, Svetlana. “Gartner’s Big Data Definition Consists of Three Parts, Not to be Confused with Three ‘V’s.” *Forbes*. March 27, 2013.
<https://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/>

Siegler, M.G. “Eric Schmidt: Every 2 Days We Create As Much Information As We Did Up To 2003.” *Tech Crunch*. August 4, 2010. <https://techcrunch.com/2010/08/04/schmidt-data/>

Stewart, Vincent R. Lieutenant General. *Worldwide Threat Assessment : Testimony before the Armed Services Committee*, 113th Cong. February 3, 2015.
<http://www.dia.mil/News/Speeches-and-Testimonies/Article-View/Article/567087/worldwide-threat-assessment/>

The U.S. Department of Justice, Office of the Attorney General. *United States Department of Justice Fiscal Years 2014-2018 Strategic Plan*. Washington, DC, 2014.
<https://www.justice.gov/sites/default/files/jmd/legacy/2014/02/28/doj-fy-2014-2018-strategic-plan.pdf>

Title 50, *The National Intelligence Council*, U.S. Code §403-3b(b)(1).

United States Department of Defense. *Joint Intelligence*. Joint Publication 2-0. Washington, DC: Government Printing Office, October 22, 2013.
http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf.

van der Windt, Peter. “From Crowdsourcing to Crowdfunding: The Cutting Edge of Empowerment?” in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. ed. Steven Livingston and Gregor Walter-Drop. Oxford: Oxford University Press, 2014.

van der Windt, Peter and Macartan Humphreys. “Crowdfunding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time.” *Journal of Conflict Resolution*. Vol 60. Issue 4. 2016. <http://journals.sagepub.com/doi/pdf/10.1177/0022002714553104>

Warner, Gregory. “In Kenya, Using Tech to Put An ‘Invisible’ Slum on the Map.” NPR.

-
- Washington, DC. July 17, 2013.
<http://www.npr.org/sections/parallels/2013/07/17/202656235/in-kenya-using-tech-to-put-an-invisible-slum-on-the-map>
- Warner, Michael, ed., “Central Intelligence: Origin and Evolution.” *Center for the Study of Intelligence, Central Intelligence Agency*. 2001. https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Origin_and_Evolution.pdf.
- Warner, Michael and J. Kenneth McDonald. *US Intelligence Community Reform Studies Since 1947*. Washington, DC: Center for the Study of Intelligence, April 2005.
<https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/US%20Intelligence%20Community%20Reform%20Studies%20Since%201947.pdf>
- The White House. *The National Security Strategy of the United States of America*. Washington, DC, February 2015. <http://nssarchive.us/wp-content/uploads/2015/02/2015.pdf>.
- Wohlstetter, Roberta. “Cuba and Pearl Harbor: Hindsight and Foresight.” The RAND Corporation. April 1965.
https://www.rand.org/content/dam/rand/pubs/research_memoranda/2007/RM4328.pdf