# From Cheese to Fondue

## A Sensemaking Methodology for Data Acquisition, Analytics, and Visualization

Robert Spousta III, Steve Chan
Dr. Steve Chan Center for Sensemaking, Asia-Pacific Institute for Resilience and Sustainability (AIRS)
Swansea University's Network Science Research Center and Hawaii Pacific University
Swansea, Wales
spousta@mit.edu, s_chan@mit.edu

Stef van den Elzen, Jan-Kees Buenen
Eindhoven University of Technology
SynerScope BV
Helvoirt, The Netherlands
e-mail: s.j.v.d.elzen@tue.nl,
jan-kees.buenen@synerscope.com

***Abstract*—Although Big Data are being leveraged through proprietary means by a host of private enterprises for significant financial gain, there are comparably fewer examples of how to harness the power of massive data through analytics in order to enhance societal resilience and directly serve the public good. In this paper, we present a three-layer framework for conducting Collaborative Big Data Analytics, including data selection and acquisition, steps comprising the analytic process, and considerations for informative data visualization. With regard to data selection, we discuss the primary characteristics of so-called Big Data, namely the Six Vs of data Variety, Volume, Velocity, Veracity, Value, and Volatility. Next, we discuss some of the various analytical tools and techniques available for processing data, as well as methods for effectively visualizing the products of data analytics. In order to illustrate the utility of such a framework, we summarize findings from our participation in Orange Telecom's Data for Development Challenges in the Republic of Côte d'Ivoire and Senegal. We conclude that while the field of Collaborative Big Data Analytics holds great promise, the development of open-source frameworks for conducting layered analytics, combined with the continuation of data challenges, such as those recently held in West Africa, will help to generate more and better uses of the Big Data that have come to dominate our world.**

***Keywords—Collaborative Big Data Analytics; Decision Engineering; Data Visualization; Sensemaking Methodology***

## I. Introduction

Whereas the dot-com boom of the late 1990s and early 2000s ushered in a wholly novel industry, replete with information-based products and virtual services marketed via the Internet, collaborative approaches for conducting civil-centric data analytics have taken longer to develop [1]. This fact notwithstanding, the rise of the Internet of Things (IoT) has introduced unprecedented levels of artificial complexity within many cyber-physical systems, which demand constant attention, lest areas of brittleness and blind spots compromise the resilience of essential services and infrastructure that are the backbone of modern civilization. In order to adulterate this vacuity, we present a basic framework for treating data and gaining insight. This **Sensemaking Methodology** addresses three primary concerns, namely, where and how to get data, how to process and refine data into insight, and how to visualize insight in a way that supports Decision Engineering endeavors. In this manuscript, we briefly outline the system of methods that comprise our three layer framework.

The remainder of the paper is organized as follows. Section II introduces the first layer of our methodological framework; harvesting and generating data, and discusses some of the primary considerations for data selection and acquisition, including the variety of sensor platforms that are responsible for producing data. Section III presents the framework's middle layer of data analytics, and goes on to describe the basic categories of analytic tools and techniques available for data processing. Section IV addresses the framework's top layer; data visualization. Section V summarizes major findings and lessons learned from our participation in the first two Data for Development (D4D) Challenges as an exemplar of the Sensemaking Methodology for **Collaborative Big Data Analytics**. We conclude in Section VI with general thoughts on the state of the art with regard to Collaborative Big Data Analytics, and propose areas for future application of our Sensemaking Methodology.

## II. Data: Prospecting for the Gold of the Information Age

We embark on our brief journey of discovery by posing two foundational questions. First, where do data come from? And second, how do we get those data? The answers to these primary questions will guide us to an optimal data harvesting strategy, and therefore, form the base of our methodological framework. However, in order to thoroughly appreciate the complexity of these seemingly simple queries, we must first explore the basic nature of data and massive datasets. At the core, we find that the phenomenon of Big Data revolves around the "Six Vs" of Volume, Variety, Velocity, Veracity, Value, and Volatility, depicted in Table 1 below.

The Big Data phenomenon is perhaps most commonly linked with the sheer amount or Volume of data being generated by a host of remote sensors, household appliances, mobile communication devices, and human content generators worldwide that totals over 2.5 quintillion bytes of data per day [2]. Although difficult to comprehend quantitatively, these reams of data come in many forms, from the millions of photos and videos shared daily from smart phones through applications like Instagram, Snapchat, and YouTube, to raw system measurements recorded by sensors and fed into synchrophasor data concentrators and

 ISBN: 978-1-61208-423-7 

other industrial control systems [3]. In order to achieve quantitative exactitude whilst navigating complex problem sets, analysts must incorporate a maximally inclusive Variety of data types and sources. In this regard, a critical determinant in achieving perspicacity through the Sensemaking Methodology is the incorporation of diverse data. By way of example, in researching issues of infrastructural resilience, we utilize a host of data gathering mechanisms, including electric grid monitoring equipment such as Phasor Measurement Units (PMU) and Digital Fault Recorders (DFR), Unmanned Aircraft Systems (UAS), Ocean Data Acquisition Systems (ODAS), Synthetic Aperture Radar (SAR) and other weather observation tools, as well as human sensor networks in the form of crowdsourced event observation and reporting. In addition to harvesting a large variety of data, the speed with which data are generated is another equally important variable, as time-critical operations including critical infrastructure protection (CIP), emergency response, law enforcement, and national defense all must be able to sense the occurrence of anomalous events in near real-time in order to prevent loss of life and property [4]. In managing both emergency responses and routine system operations, all data consumers rely on the authenticity or Veracity of data in order to gain actionable insight. The consistency of data taxonomy is an important aspect of Veracity, and, in this regard, discovery standards for electronic resources such as the Dublin Core standards for Metadata are essential for datasets held by diverse curators to remain compatible with one another [5].

TABLE 1 CHARACTERISTICS OF DATA

| V | The 6 Vs of Big Data | |
|---|---|---|
| | ***Description*** | ***Units of measure / Dimensions*** |
| ***Volume*** | Massive amounts of data | Bytes => Terabytes |
| ***Variety*** | Multiple forms / formats | video, sms, .pdf, .doc, .jpg, .xls, .rtf, .tif, PMU, etc |
| ***Velocity*** | Speed of data feeds | Event-driven / Streaming |
| ***Veracity*** | Trustworthiness of data | Provenance / Pedigree |
| ***Value*** | Usefulness of data | Ambiguity / Uncertainty; Correlation / Causation |
| ***Volatility*** | Shelf-life of data | Time-Sensitive / Static |

a. An alternate V of Viability has also been proposed in [2], which we believe is subsumed above

A more persistent challenge for data Veracity is the ability to establish the provenance and pedigree of data, particularly in the context of data manipulation and spoofing, or counterfeiting in the information supply chain. While gathering redundant data from multiple sources, and cross-referencing particularly specious data are prudent strategies for mitigating the negative impact of false or corrupted data, ensuring data Veracity is a perennial problem that demands consistent attention and focus.

Two rather more subjective aspects of data are their Value and Volatility. In **Decision Engineering**, the Value of a given dataset loosely correlates to how much of any given decision can be built from it. In other words, can we decide a course of action based on a single dataset? If so, then that dataset could be said to be of high Value. If many disparate datasets are required in order to engineer a single decision, then each of those datasets is of comparatively low Value. Data’s Volatility or duration of relevance depends largely on the nature of the decision it is serving to inform or build. Whereas certain digitally preserved historical records maintain their relevance or Value in perpetuity, other datasets that pertain to rapidly evolving circumstances may remain relevant for only a matter of days, if not seconds. Determining a dataset’s Volatility is yet another important step in the process of Sensemaking.

Having established the basic nature of data, we return to the original question of where and how to acquire data. For all organizations - public, private, and any permutation in between - data accessibility and knowledge management remain areas of active research and constant improvement [6]. With the United Nations (UN) recently asserting that information in itself is a life-saving need for people in crisis, just as important as water, food, and shelter, the necessity of open source data is clearly a global one that now transcends the realm of scholarly open access [7]. So, the short answer to our question is that there is no comprehensive, authoritative single source for all data, and therefore, we get data from everywhere we can, however we can.

## III. STACKING THE DECK: TOOLS AND TECHNIQUES FOR LAYERED DATA ANALYTICS

Next, we turn to the analytic component of the Sensemaking process, which includes algorithms, cognitive high performance computing, machine learning, signal resolution, allegorical engines, and the Unstructured Information Management Architecture (UIMA). Many of these components are rooted in mathematical concepts dating back centuries. Notable examples include the famous problem of the Bridges of Konigsburg and Graph Theory, Ada Lovelace’s development of early programming instructions for Babbage’s Decision Engine, the Pragmatists’ precepts of indeterminacy, order in chaos, and long-run convergence; as well as Turing’s Machine, and Weaver’s Complex Systems Ontology [8].

The modern analytical toolkit is comprised of far too many instruments to concisely summarize here. However, there are fundamental components of the analytic process, which we will introduce in this manuscript. Upon identifying, generating, and acquiring data, the initial step in the analytic layer of our framework is data ingestion and refinement. By way of example, satellite imagery is unfortunately not as simple as an “eye in the sky” beaming down neat pictures to a computer console for analysis and distribution. The many 0’s and 1’s that make up the digital representation of a physical object must first be processed and translated into an intelligible picture. Once raw data are refined into a malleable commodity, that commodity can then be annealed into meaningful insight through a systematic layering of Analytics on Analytics (A2O). This process begins with a geospatial and or temporal matrix of data points, and proceeds through a set of systematic organizational steps that include data clamping, normalization, and hierarchical clustering, in order to reveal traces of emergent phenomenon and achieve pattern recognition. Such patterns are the bedrock of insight, and serve to evaluate the role of myriad variables in the emergent outcomes of complex systems and networks, as depicted below in Figure 1. However, a fundamental prerequisite for effective A2O is the storage and management of massive datasets. In this regard, distributed computing architectures and parallel processing are also prominent features in the analytic layer of the Sensemaking Methodology [9].

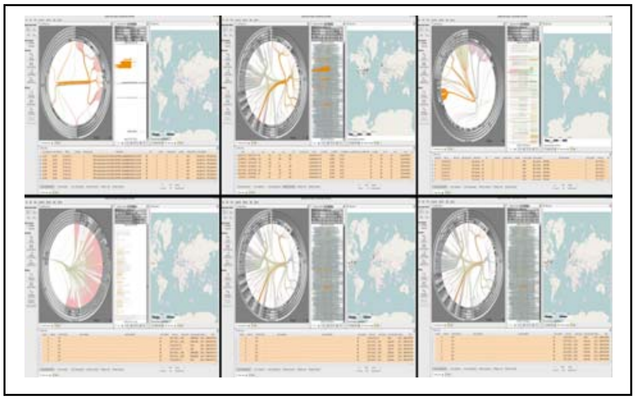

Figure 1. Example of SynerScope A2O Visualization Suite

Impressive though they may be, machine capabilities comprise but one half of the analytic layer of our methodological framework. The remaining half relies on the inherently human capabilities of contextual orientation and intuitive leaping [10]. Whereas machines are capable of generating, processing, and storing massive quantities of data, the human mind remains unique in its ability to superimpose context over data in order to discern relevance and meaning. Hence, the Sensemaking Methodology is characterized by its counterpoising and fusion of socio and techno perspectives. On the one hand, we leverage the technical advantages of machine capability to yield algorithmic insight. On the other hand, we also leverage inherent knowledge of the human social condition and sentient thought to arrive at heuristic insight. This socio-techno unification is at the heart of our methodology for pattern recognition and Decision Engineering. Going back to the example of satellite imagery, let us consider the case of the Global Earth Observing System of Systems (GEOSS) and the view of Somali villages at night as an illustration of counterpoising algorithmic versus heuristic insight. With the rise of both maritime piracy off the coast of the Horn of Africa, and the violent extremist organization Al-Shabaab in Somalia, international security organizations were keen to establish a link between the two groups [11]. As assets in the GEOSS satellite constellation observed significant variances in the night-time illumination of various towns along the Somali Coast and provincial capitals, analysts sought to employ the algorithmic insight as evidence for a correlation between the dispensation of pirate ransoms and the buildup of jihadi strongholds [12]. However, heuristic insight suggested that the ideological and religiously-motivated nature of Al-Shabaab was incompatible with the financially-driven motives of the criminal piracy network, and therefore a link was unlikely. The truth of this insight would later be established through data gathered by the International Criminal Police Organization (INTERPOL) and the United Nations Office on Drugs and Crime (UNODC) [13]. Such an example shows us that while technology and algorithmics are more than capable of identifying patterns of interest, we still need heuristic insight to decipher what those patterns actually mean.

## IV. A PICTURE TELLS A THOUSAND WORDS: IMPARTING INSIGHT THROUGH DATA VISUALIZATION

Upon recognizing patterns of interest, we are now ready to move into the third and final phase in the **Sensemaking Process**; visualizing insights for Decision Engineering. The primary aim of the data visualization phase is to establish the relevance of insight gained through the A20 process, and ultimately answer the basic question of “So what?” Figure 1, above, displays output from one of our visualization platforms, the SynerScope. SynerScope and other similar tools use a coordinated multi-view approach with a scalable and flexible visual matrix in order to visualize key insights from massive datasets.

However, before we progress into any further detail with regard to contemporary visualization techniques, let us briefly consider the history of data visualization. The roots of visualization are as old as human knowledge and communication; from cave paintings, to pictographs, hieroglyphics, numerology, symbolic logic, and language. In order to understand what methods have been developed over time for effectively conveying knowledge and information, it is instructive to visit certain historical examples. One case in point is the work of the Mixtec civilization of Oaxaca, Mexico [14], depicted below in Figure 2.

Figure 2. Image from the *Codex Vindobonensis Mexicanus*

Although the figure above depicts the Mixtec's primordial cosmology and creation mythology, it is an early example of how human insights gained through observation of natural phenomenon (i.e., data analysis) were preserved for distribution and posterity. This and other similar precedents from early civilization remain germane to many data-related fields, including Education, the Arts, Public Information, Manufacturing, Product Advertisement, Device Instruction Manuals, Traffic Signage, Emergency Management, and Information Technology (IT) [15]. With the advent of the Internet, and eventually the World Wide Web, the tradition of data visualization has continued to evolve. Today, such professional disciplines as Cognitive Science, Behavioral Psychology, Computer-Assisted Design (CAD), and Strategic Communication all build on the work of early visualization specialists by combining machine capability with human insight to generate socio-techno innovations in how the brain senses and interprets information. In turn, our interpretation and assimilation of information drives our ability to engineer decisions and determine appropriate courses of action, as individuals in daily life, as agents in organizations, and as members of the global citizenry.

Nevertheless, this does not mean that modern data visualization is a perfected science. Rather, visualization is a principled art that requires both intelligence and intuition in its composition. In turn, efforts to visualize pseudo-insights that are not informed by robust A2O run the risk of proliferating misinformation, bias, conflict, and spoilage of resources [16]. In addition to these pitfalls, data-informed visualizations also can be subject to information overload, if insights are not concisely crystallized in a digestible form, as depicted in Figure 3 [17].

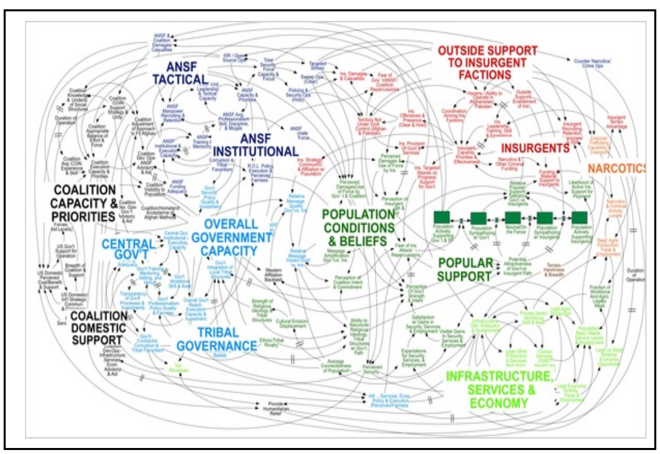


Figure 3. Example of Counterinsurgency Diagram

The design of any given data visualization is driven by two primary factors; the nature of the decision it serves to engineer, and the demographic characteristics of the audience or consumer. Firstly, is the aim of the visualization simply to impart generally useful information, or is it intended to inform a specific choice? If the aim is the former, then visualizations such as that in Figure 3 may be appropriate. However, decision-quality visualizations must clearly depict actionable intelligence, and offer tangible courses of action. Secondly, how much does the target audience for a given data visualization already know? An audience of laymen will require a significant amount of context in order to make sense out of visualizations. Conversely, too much context will be superfluous (and potentially distracting) to an audience of experts. Therefore, constructing an effective data visualization means striking a delicate balance between sufficient context and specific insight.

With this in mind, we turn to a final consideration regarding the value of data visualization; the identification of brittleness in complex systems. In light of the staggering layers of complexity and interdependence that characterize many of our most critical infrastructural systems (e.g., electric grids, the Internet, etc.), there is significant potential for percolation effects or cascading failure [18]. Therefore, to ensure the resilience of such systems, it is essential to identify areas of brittleness or weak links in the chain before they fail. With regard to the resilience of the Internet in particular, tools such as the SeeSoft System, pictured below in Figure 4, enable analysts to visualize statistics of interest in software code [19]. In the case of Figure 4, a color-coding scheme displays how recently lines of code have been changed, with red lines having been most recently changed, and green lines having remained unchanged the longest.

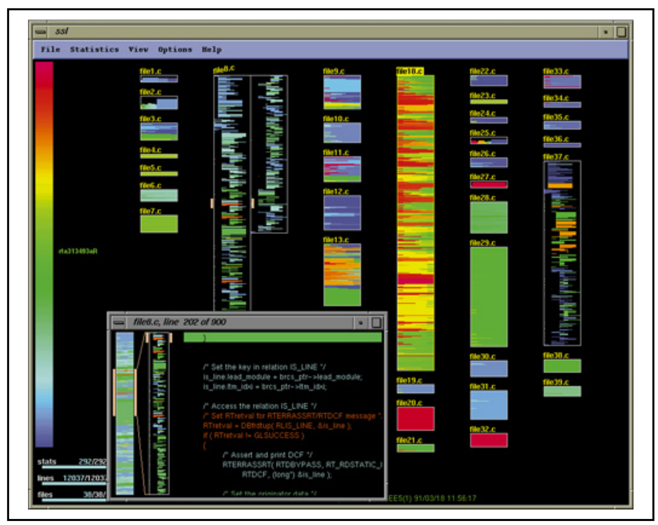


Figure 4. SeeSoft software code visualization system, Lucent Technologies

Visualization tools are invaluable assets that enable us to quickly and clearly see areas of potential brittleness in complex systems. In the case of Figure 4, above, we have a mechanism to visualize answers to questions such as whether software security improves with age, as lines of code not recently updated to address proliferating cyber threat vectors are likely brittle [20]. Therefore, visualization is not only a product of the analytic phase of the

Sensemaking Methodology, but can actually be a feedback loop that helps to inform the A2O process.

## V. PROOFS OF CONCEPT: SYNERSCOPE AND THE DATA FOR DEVELOPMENT CHALLENGE

With our **Sensemaking Methodology** in hand, we finally come to the shores of West Africa and the Data for Development Challenge (D4D) [21]. Since its inauguration in 2012, the annual D4D Challenge has represented a unique opportunity for Big Data analysts to experiment with diverse tools and techniques for harvesting insight from mobile phone data. For each challenge, international competitors from academia and private industry are given the chance to analyze a multitude of datasets pertaining to mobile phone use in a designated country during a circumscribed portion of the year [22]. We have had the privilege to participate in both challenges thus far, in the Republic of Côte d'Ivoire and Senegal, with a sampling of our results displayed below in Figure 4 [23].

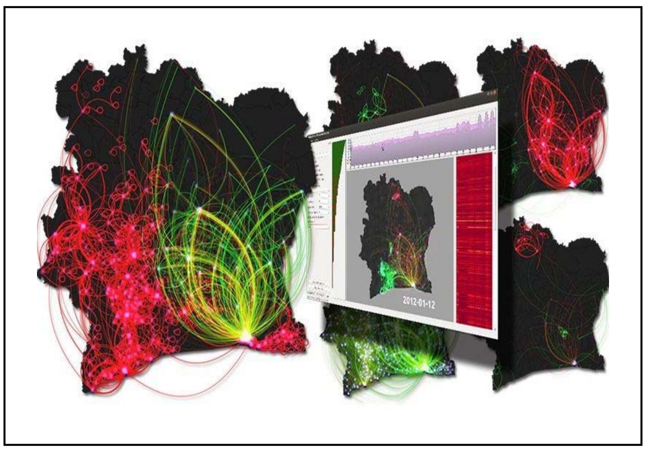

Figure 4. 2013 D4D Best Visualization prize winner: *"Exploration and Analysis of Massive Mobile Phone Data: A Layered Visual Analytics Approach"*

In conducting our analysis of the D4D datasets and generating the illustrations sampled above, two lessons became clear to us. First, we needed data Variety, through which to contrast and correlate mobile phone activity with other significant trends and events. For the first D4D in Côte d'Ivoire, we contrasted the given mobile phone data with UN reports of violent conflict and significant social disturbance, as well as meteorological data for the given timeframe. This helped to reveal regional political affiliations and ethnic enclaves, as violent events targeting certain political and ethnic groups in the capital city, Abidjan, catalyzed notable increases in call activity to specific communities elsewhere in the country. In addition, we observed that abundant rainfall in areas of significant cocoa and yam cultivation correlated with heightened call activity, likely indicating increased agro-business developments at specific points in the growth and harvest cycles in response to favorable weather conditions. Our second lesson learned was the need to adopt multiple perspectives from which to interrogate the datasets. Our normalization and clustering algorithms produced dendograms, with which we were able to sort items (e.g., cell towers) of similar behavior into groups for further investigation. By grouping cell towers of similar call behavior, we were then able to further explore what other commonalities linked these disparate regions.

Although such techniques are still relatively nascent, we believe that the work of our team and fellow D4D participants is a clear demonstration that Collaborative Big Data Analytics can help to increase insight into complex interrelated phenomenon, and thus improve Decision Engineering in a variety of social, political, and economic arenas. However, the implementation of our **Sensemaking Methodology** remains in the early stages, and inevitably there is room for improvement in such an approach. Specifically, increasing the Volume and Variety of data included in the A2O phase will yield greater insight in future D4D Challenges, and other applications of our methodological framework. In addition, the deliberate articulation of alternate frameworks for Collaborative Big Data Analytics will help to progress the state of the art, by revealing common best practices as well as shortfalls and gaps.

## VI. CONCLUSION: STANDING ON THE THRESHOLD OF A BRAVE NEW WORLD

Our journey ends with the realization that humanity's quest for insight is by nature eternal. Although it is temporally little, the story of Big Data is truly epic. As machine capability continues to accelerate, the power and promise of data analytics will only grow. At the same time, our ability to make sense out of evolving circumstances quickly, and adapt social structures accordingly will be important determinants in the shape of things to come.

Our experiences with D4D and other instances of Collaborative Big Data Analytics are evidence that critical thinking is an inseparable ingredient in the recipe for Big Insight, and that socio-techno approaches are an indispensable element of complex problem solving. We believe that open and inclusive approaches such as the **Sensemaking Methodology** have the potential to enhance numerous dimensions of resilience, including those of cyber-physical systems, societies, and individuals. Systematic Decision Engineering is a practical way to identify latent Black Swan blind spots, Maginot Line-scale brittleness, and Pearl Harbor-level threat vectors. Similarly, we also hope that such a methodology can facilitate positive developments, such as the smart integration of green technologies into sustainable Blue Economies [24], and an improvement in our roles as both environmental stewards and engines of social progress. Each of these areas represents exciting and relatively unexplored realms of

 ISBN: 978-1-61208-423-7 

research that we have designated as targets for future work. Specifically, we plan to demonstrate how technological advancements such as Pervasive Remote Sensing (PRS), Comprehensive Domain Awareness (CDA), and Cognitive Computing can be effectively integrated with human Sensemaking techniques to achieve increasingly useful insights and practical Decision Engineering solutions.

## ACKNOWLEDGMENT

The authors would like to thank the Cyber Futures Center, an initiative of the Sensemaking-U.S. Pacific Command Fellowship, and the Dr. Steve Chan Center for Sensemaking — one of the centers of the Asia-Pacific Institute for Resilience and Sustainability (AIRS), which is jointly anchored at Swansea University's Network Science Research Center and Hawaii Pacific University — for the opportunity to study the challenges facing Hawaii and other archipelagos, and to contribute towards the various Public Private Partnership Initiatives aimed at developing solutions to overcome those challenges.

## REFERENCES

[1] N. R. Council, *Frontiers in Massive Data Analysis*. Washington, DC: The National Academies Press, 2013.

[2] N. Biehn. (2013, May) The Missing V's in Big Data: Viability and Value. *Wired*. Available: http://www.wired.com/2013/05/the-missing-vs-in-big-data-viability-and-value/ accessed May 20, 2015

[3] C. Alcaraz and J. Lopez, "Wide-Area Situational Awareness for Critical Infrastructure Protection," *Computer,* vol. 46, pp. 30-37, 2013.

[4] K. M. Chandy, "Sense and respond systems," in *Int. CMG Conference*, 2005, pp. 59-66.

[5] S. Weibel, J. Kunze, C. Lagoze, and M. Wolf, "Dublin core metadata for resource discovery," *Internet Engineering Task Force RFC,* vol. 2413, p. 132, 1998.

[6] M. Alavi and D. E. Leidner, "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues," *MIS Quarterly,* vol. 25, pp. 107-136, 2001.

[7] C. Hajjem, S. Harnad, and Y. Gingras, "Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact," *arXiv preprint cs/0606079,* 2006.

[8] W. Weaver, "Science and Complexity," *American Scientist,* vol. 36, pp. 536-544, 1948.

[9] G. S. Sureshrao and H. P. Ambulgekar, "MapReduce-based warehouse systems: A survey," in *Advances in Engineering and Technology Research (ICAETR), 2014 International Conference on*, 2014, pp. 1-8.

[10] N. Ford, "Information retrieval and creativity: towards support for the original thinker," *Journal of Documentation,* vol. 55, pp. 528-542, 1999.

[11] J. Stevenson, "Jihad and Piracy in Somalia," *Survival,* vol. 52, pp. 27-38, 2010/03/01 2010.

[12] A. Shortland, "Treasure mapped: using satellite imagery to track the developmental effects of Somali Piracy," *London: Chatham House,* 2012.

[13] S. Yikona, *Pirate Trails: Tracking the Illicit Financial Flows from Pirate Activities Off the Horn of Africa*: World Bank Publications, 2013.

[14] B. E. Byland and J. M. Pohl, *In the realm of 8 Deer: The archaeology of the Mixtec codices*: University of Oklahoma Press, 1994.

[15] J. Z. Gao, L. Prakash, and R. Jagatesan, "Understanding 2D-BarCode Technology and Applications in M-Commerce - Design and Implementation of A 2D Barcode Processing Solution," in *Computer Software and Applications Conference, 2007. COMPSAC 2007. 31st Annual International*, 2007, pp. 49-56.

[16] W. Neil Adger, N. W. Arnell, and E. L. Tompkins, "Successful adaptation to climate change across scales," *Global Environmental Change,* vol. 15, pp. 77-86, 2005.

[17] E. Bumiller. (2010, April 26) We Have Met the Enemy and He is Powerpoint. *New York Times*. Available: http://www.nytimes.com/2010/04/27/world/27powerpoint.html?_r=2 accessed May 20, 2015

[18] S. H. Strogatz, "Exploring complex networks," *Nature,* vol. 410, pp. 268-276, 2001.

[19] S. G. Eick, J. L. Steffen, and E. E. Sumner, Jr., "Seesoft-a tool for visualizing line oriented software statistics," *Software Engineering, IEEE Transactions on,* vol. 18, pp. 957-968, 1992.

[20] A. Ozment and S. E. Schechter, "Milk or wine: does software security improve with age?," in *Proceedings of the 15th conference on USENIX Security Symposium-Volume 15*, 2006, p. 7.

[21] J. K. Laurila, D. Gatica-Perez, I. Aad, J. Blom, O. Bornet, T. M. T. Do, *et al.*, "From big smartphone data to worldwide research: The Mobile Data Challenge," *Pervasive and Mobile Computing,* vol. 9, pp. 752-771, 2013.

[22] V. D. Blondel, M. Esch, C. Chan, F. Clérot, P. Deville, E. Huens, *et al.*, "Data for development: the d4d challenge on mobile phone data," *arXiv preprint arXiv:1210.0137,* 2012.

[23] J. Poole. (2013, May 6) Winning Research from the Data 4 Development Challenge. *United Nations Global Pulse*. Available: http://www.unglobalpulse.org/D4D-Winning-Research accessed May 20, 2015

[24] G. Pauli, "The blue economy," *Our planet,* pp. 24-27, 2010.

 ISBN: 978-1-61208-423-7 