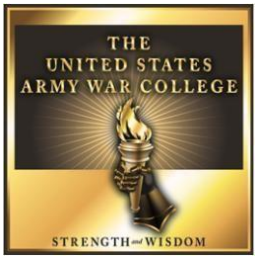
Strategy Research Project

A Clearer Picture: Using Big Data to Better Understand Megacities

by

Lieutenant Colonel Robert G. Dixon
United States Army



United States Army War College
Class of 2015

REPORT DOCUMENTATION PAGE

Form Approved--OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-04-2015	STRATEGY RESEARCH PROJECT	

4. TITLE AND SUBTITLE
A Clearer Picture: Using Big Data to Better Understand Megacities

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Lieutenant Colonel Robert G. Dixon
United States Army

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Dr. Steve Metz
Strategic Studies Institute

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army War College, 122 Forbes Avenue, Carlisle, PA 17013

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A: Approved for Public Release. Distribution is Unlimited.

13. SUPPLEMENTARY NOTES
Word Count: 7834

14. ABSTRACT

Global urbanization has given rise to the megacity: large, densely-packed urban areas with more than ten million people. These cities are becoming increasingly connected, dense and complex, and are becoming non-state actors' preferred hiding places. It is inevitable that the US military will find the need to operate in this environment, and it is currently unprepared to do so. Gaining and maintaining a comprehensive understanding of the environment is the first challenge to preparing forces for megacity urban operations. The emerging fields of big data, analytics, and data-driven decision-making offer significant potential towards this end. The Army must study this environment and the emerging sets of tools if it wishes to remain relevant in tomorrow's fight.

15. SUBJECT TERMS
Megacity, Design, Urban Operations, Analytics

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(w/ area code)*
UU	UU	UU	UU	41	

USAWC STRATEGY RESEARCH PROJECT

A Clearer Picture: Using Big Data to Better Understand Megacities

by

Lieutenant Colonel Robert G. Dixon
United States Army

Dr. Steve Metz
Strategic Studies Institute
Project Adviser

This manuscript is submitted in partial fulfillment of the requirements of the Master of Strategic Studies Degree. The U.S. Army War College is accredited by the Commission on Higher Education of the Middle States Association of Colleges and Schools, 3624 Market Street, Philadelphia, PA 19104, (215) 662-5606. The Commission on Higher Education is an institutional accrediting agency recognized by the U.S. Secretary of Education and the Council for Higher Education Accreditation.

U.S. Army War College
CARLISLE BARRACKS, PENNSYLVANIA 17013

Abstract

Title:	A Clearer Picture: Using Big Data to Better Understand Megacities
Report Date:	01 April 2015
Page Count:	41
Word Count:	7834
Key Terms:	Megacity, Design, Urban Operations, Analytics
Classification:	Unclassified

Global urbanization has given rise to the megacity: large, densely-packed urban areas with more than ten million people. These cities are becoming increasingly connected, dense and complex, and are becoming non-state actors' preferred hiding places. It is inevitable that the US military will find the need to operate in this environment, and it is currently unprepared to do so. Gaining and maintaining a comprehensive understanding of the environment is the first challenge to preparing forces for megacity urban operations. The emerging fields of big data, analytics, and data-driven decision-making offer significant potential towards this end. The Army must study this environment and the emerging sets of tools if it wishes to remain relevant in tomorrow's fight.

A Clearer Picture: Using Big Data to Better Understand Megacities

> The adversary is occupying the cities. Operations in exceptionally large urban environments is an enormous challenge… and is an environment that the Army is currently ill-prepared to successfully operate within.
>
> —Chief of Staff of the Army's Strategic Studies Group
> Megacity Concept Team[1]

The Army and the Megacity Environment

As the Army peers into the future to prepare for what it may be called upon to do, it is obvious that the world of tomorrow will be increasingly urban. The number and size of large urban areas continues to grow, and are quickly becoming the dominant form of human habitation. The scale, connectedness, density, context and complexity of these environments make them the greatest challenge to military operations.[2] The US Army is unprepared for operations in a modern megacity. This unpreparedness is evident in multiple categories, including equipment, munitions, doctrine, and training. Most importantly, the US Army lacks a coherent intellectual approach for operating in large urban areas. Despite doctrine's call to "think systemically" about cities, all of the doctrinal methods of analysis are reductionist algorithms, approaching cities as if they are reducible structures that can be dealt with "one bite at a time."[3] The Army needs a better way to achieve holistic understanding of cities. "Big data"[4] and wiki-sourcing[5] can help the Army see the whole city in addition to its parts if it learns to exploit them. These tools have the ability to provide commanders with a deeper, more sophisticated understanding of the urban ecology—far beyond what was previously possible.

US Military Joint Doctrine suggests multiple challenges to operating in this environment: cities may reduce the advantages of technology, are manpower intensive, decentralized, time-consuming, produce large numbers of casualties (many civilian),

restrict Rules of Engagement, and impact weapons employment and effects.[6] Additionally, the ubiquitous presence of media outlets and the democratization of information (e.g. social media) impacts every aspect of urban operations. Yet, Joint Doctrine still says, ". . . ground combat [in urban areas] may be the most effective and efficient way for a commander to accomplish operational or strategic objectives."[7] The credibility of this doctrinal statement has been tested in multiple operations from Baghdad to Hue City to Aachen. In reality, ground *combat* in urban areas is often the least effective and efficient way to achieve operational or strategic objectives. Despite military preference to avoid operating in cities, they are sometimes unavoidable. As cities increase in size and complexity, they become more attractive to adversaries who seek to use these environments to counter US firepower and precision superiority. Further, the Army cannot be content to operate in smaller cities with immature infrastructure: it must be prepared to operate in even the largest, most modern cities. The Army will not always have the option to choose where it conducts operations. If the Army intends to take the fight to the enemy, then it must prepare to operate in megacities.

Doctrine provides some guidance on the way forward in the megacity environment, suggesting that these complex settings require "different approaches."[8] Other US or international government, nongovernmental and private organizations will impact military operations; cities must be considered in their context of suburban and rural areas; and the heterogeneity of cities mean that solutions and challenges in one part of the city may be completely different in other parts of the same city.[9] These considerations are necessary but not sufficient for appreciating the complexities of the

environment. Cities are the most complex operating environment a land force may operate within. Not only must the force deal with an adversary and extremely complex terrain, the city itself sustains millions of lives and is connected to billions of others.

The greatest challenge is understanding the environment itself—identifying those aspects that are critical to both mission accomplishment and to the survival of the city. To achieve this understanding, being able to detect deviations from “normal” urban patterns is an “invaluable capability.”[10] As the Joint Publication for urban operations states, “Understanding urban patterns supports assessment and has value at the strategic, operational, and tactical levels. Understanding urban patterns is a time-consuming effort and requires patience, attention, and the focused observation skills of the joint force.”[11]

This last assertion must be challenged. Achieving a robust, sophisticated understanding doesn’t need to be time consuming, and it doesn’t need to take the whole joint force: millions of sensors already exists within the environment, and machine-learning advanced analytics can reveal more information than an army of analysts working around the clock. The military must change the way it thinks about collection and analysis. The use of big data and wiki-collaboration can lead to innovative, relatively inexpensive and rapid ways to achieve a more accurate understanding of the urban environment.

The Relevance of Megacities

More than half of the world’s population currently lives in urban environments, and experts expect that more than 66% will be urban by 2050.[12] With the increase in urban dwellers comes a significant increase in the size of the cities. For example, in 1990 there were only 10 cities designated as “megacities” (cities with a population

exceeding 10 million). In 2014, there were 28 megacities, and 14 more are expected to emerge by 2030.[13] Urban areas are growing rapidly and are quickly becoming the most important features in nearly every country, dominating the social, economic and political spectrums at the near-total exclusion of rural areas.

Rapid growth is fuelled largely by rural migrants, and consists increasingly of relatively young populations. United Nations–Habitat expects that by 2030, 60% of all urban inhabitants will be under the age of 18, with a majority living in slums or informal settlements.[14] Most of this rapid growth occurs in the developing world, in cities that have the least capacity to integrate the new population. The resulting growth spawns rapid increases in slums, informal economies, and widening separation between the rich and the poor. As much of the rapid growth in developing world cities occurs in coastal cities, rising sea levels are yet another factor that can dramatically increase the scale of unrest or disaster. These features, along with the interconnectedness of modern cities make many of them a strategic concern for the United States.

The Army is hesitant to consider large urban environments as productive operating environments. The age-old advice of military theorists has been to avoid urban combat whenever possible. Yet as a significantly higher percentage of the world's population urbanizes, it is unrealistic that the military's leaders can turn to the National Command Authority and tell them that they have no options available—short of long distance bombing—for affecting a strategic problem in a megacity. As Col. Kevin Felix, Chief of the Future Warfare Division at the Army Capabilities Integration Center stated, "The Army is about creating options for the president and the national command authority." The Army must be able to offer "multiple options related to multiple dilemmas

across multiple domains," from humanitarian assistance to combat operations.[15] These dynamics suggest that the military must be prepared to operate in and around large urban areas. The probability is that urban operations in a massive urban area are unavoidable and inevitable.[16]

The military currently lacks much of what is needed to operate successfully in the future urban environment. Equipment, training, doctrine, non-lethal capabilities, etc. are all areas where shortfalls exist. Yet, given the scale of the modern megacity, even massive numbers of the best equipment will be insufficient to dominate the environment. Instead, the Army must learn to see cities differently, and must develop approaches that allow it to operate within the urban environment without the need to control or destroy it.[17]

Improving the Ability to See the City

There are numerous areas that can be investigated to increase the Army's ability to understand and operate in the future urban environment.[18] Yet, given the characteristics of the megacity, there is little chance of any force dominating the environment as militaries have sought to do in the past. Instead of a narrow ecological view, simply seeing terrain with "human terrain" placed on top of it, leaders must begin to comprehend the nuances of a complex ecosystem.[19] Current doctrinal methods of seeing the city are slow, labor intensive, and have not yielded quantifiable advantages to the warfighter. While the goal of applying social science to the operational environment (as the Human Terrain Teams do) is laudable, big data analytics offers dramatic improvements on outcomes. The greatest breakthroughs to understanding urban ecologies will likely be made through expanding the ability to see physical, economic, social and political changes in the city in real time.

The City as an Organism

Many urban thinkers use the metaphor of a living organism to describe a city.[20] Both have sub-systems that have specific functions that work together to produce behavior that is more than just the sum of its parts. Cities, like organisms, take in energy and expel waste. Cut off from its energy source, the organism and the city starve; inhibit the removal of waste and they become toxic.

The interaction of the subcomponents of an organism combine in unpredictable ways. While rigorous study of biology has led to significant improvements in understanding organisms, a complete understanding of all functions, and the higher order effects of sub-system failure in the human body still evades the medical community.[21] This presents a useful metaphor.

Despite years of focus on one type of organism (i.e. humans), medical doctors still approach each patient as unique. Most doctors do not fixate on the injury itself, but instead investigate the body as a whole, collecting a thorough patient history, bloodwork, and vital statistics before conducting an operation. Long-term monitoring is the basis for informed surgery, and instrumentation and monitoring allow the doctor to observe the health of the whole person, even when they are working on an isolated injury. Constant monitoring of vital statistics during an operation gives the surgeon early indicators that an action in one part of the body is having a negative effect on another part. Doctors also prioritize the survival of the patient and the functioning of the vital systems over treatment of a specific malady. When treating a traumatic injury, for example, maintaining airway and circulatory systems take priority over treating the specific wound, even if it means removing a limb. In short, keeping the patient alive and understanding the systemic effects of an operation are primary concerns.

Doctors do not ignore the similarities between patients (they are all human after all), but instead focus investigation on identifying differences (blood type, history, environmental factors, etc.) to determine the appropriate courses of treatment. Without specific knowledge and constant monitoring (i.e. a holistic view of the patient) the physician could do more harm than good.

If the metaphor holds, current doctrinal approaches to urban operations are more akin to 19th Century medical practice (where amputation and bloodletting were common) than to the 21st Century approach. The Army's approach to operating in cities often causes severe damage to the physical and social structure of the city, leading to a weakened organism that is susceptible to infection and decay. Without a thorough understanding and monitoring of the city's vital statistics, operations can often do more harm than good.

Understanding Flow Systems

Applying the living organism metaphor, flow systems act as the arteries, digestive system and sensory organs of the city. Information, money, food, water, electricity, people—everything that moves generates a system that evolves towards greater efficiency.[22] Monitoring these systems can reveal much of what is hidden in the city, much like lab tests reveal hidden clues to a person's health. This is not novel: alert commanders have investigated anomalies in their areas of operations that have revealed hidden changes.[23] What is new is the quality and timeliness of the information that is available through modern sensors, and the enormous quantity of data that is produced in the modern urban environment.

The Army currently lacks the resources, expertise or approaches to investigate and exploit this reservoir of information. The use of existing data streams and the ability

to use them to dynamically map a city's flow systems can give operators an unprecedented understanding of the environment. Real-time aggregate mobile phone data, utility usage, traffic flows, and cyber activity all reveal changes in the city's flows and give operators vital information about what is happening within it. This information is useful in two important ways: it exposes the impact operations have on the city, and it reveals the manner in which the city itself is responding to the operations. This can be in the form of illuminating an adversary's reaction, or simply a change in the population's movement and actions that indicates what types of future actions might take place. Achieving this level of clarity requires large amounts of data, but in most places much of the data already exists. Governments and major corporations (even within the poorest cities) collect billions of data points a day, and the quality and volume is increasing exponentially.

How Doctrine Currently "Sees the City"

Joint Intelligence Preparation of the Operational Environment (JIPOE) is the analytic process joint intelligence organizations use to produce intelligence assessments and estimates.[24] JIPOE is a reductionist process that categorizes information into "physical areas and factors" and the "informational environment." The informational environment is the aggregate of individuals, organizations and systems that collect, process, disseminate, or act on information, and it is comprised of the physical, informational and cognitive dimensions.[25] The operational planning counterpart to JIPOE is Intelligence Preparation of the Battlespace (IPB). It is formally defined as "the systematic, continuous process of analyzing the threat and environment in a specific geographic area." IPB is a more tactical tool requiring a higher degree of fidelity.

It follows four steps: define the battlefield area, describe the battlefield's effects, evaluate the threat, and determine adversary courses of action.[26]

These steps are fine as far as they go, but the processes are fixated on the "battlefield" and the enemy or adversary, and spend little effort truly investigating the ecology of the urban setting. The level of understanding required to operate in a complex, modern city far exceeds that of the past operating environments.

Current Army doctrine, for example, does not consider how the city itself can provide much of the necessary information. Rather, Army doctrine expects units to expend considerable energy attempting to collect data through traditional means. Army Techniques Publication 2-01 (ATP 2-01), *Plan Requirements and Assess Collection*, provides lengthy lists of the kinds of information that can be garnered from different kinds of military units and nonmilitary sources for use in urban operations.[27] All of the sources listed require intelligence collectors to meet with people (presumably in person) to garner their *opinion* or (in some cases) collect physical maps or plans from them. The manual also notes that, in the case of nonmilitary officials, "These authorities may be biased . . . and their advice is almost certainly slanted in their favor."[28] There is no mention of exploiting existing data sets to garner the same information, most of which is available without risking the bias or ineptitude of opinions. ATP 2-01 further prescribes which kinds of intelligence could provide the most useful information for urban operations. Among the eight intelligence considerations listed in the manual, only open-source intelligence mentions databases or electronic media. Open-source intelligence, however, is limited to openly-provided or publicly available information, limiting its utility in developing a comprehensive understanding when compared to aggregate big data.[29]

Additionally, as each intelligence discipline is based almost exclusively on requirements generated by Combatant Commanders or their service components, there is currently no intelligence focus on urban environments as a unit of analysis. Instead, most intelligence is focused on specific adversaries.

But the demands of achieving strategic goals in the future will increasingly rely on the military's ability to develop a deep understanding of the urban ecology. Limiting investigation of the environment to a well-defined operational area, or fixating inquiries only on the adversary automatically limits the options the Joint Force commander might consider. A thorough understanding of context (including the historical, social, political, cultural, and economic dynamics) in addition to a robust appreciation for the scale, density, connectedness, complexity and flow systems may reveal novel solutions to complex problems. Just as many cities are beginning to use big data to manage daily operations, military leaders can tap into these same data networks that offer predictive analysis and currently reveal and track citizen sentiment, emergency response times and performance, utility efficiency and repair, and traffic flows.[30]

The doctrinal approach to understanding urban areas treats cities as if they were closed systems, often ignoring the myriad connections with the rest of the world. This problem is exacerbated in places where a majority of external connections link a city to another geographic Combatant Command's Area of Responsibility such as Lagos, Tokyo, and Dubai which are global cities.[31] Additionally, doctrinal planning approaches to cities treat them as mechanical entities, where linear cause and effect models work in expected, predictable ways.[32] The result of these approaches is too often a linear, deterministic approach with clearly defined goals and end-states that have little hope of

success. Despite the uniqueness of the environment, doctrine does not entertain changing the intellectual approach, instead applying a process that produces a limited number of options.[33]

Doctrine instructs leaders to seek out civil leaders to “handle noncombatants.” This may work in areas of the city where civil leaders are actually the dominant authority. More often, there are large areas of the city where alternative governance structures dominate. In many of the favelas of Brazil, for example, drug lords provide governance instead of any central authority. The same can be said of the “mayors” of slums in Lagos, the Basti in Dhaka or any other large urban area. This dynamic is not limited to the “developing world.” Large parts of some European or US cities have struck uneasy alliances with alternate control structures, evidenced by some of the immigrant neighborhoods where city officials are unwilling or unable to enforce their laws.[34]

It is often fashionable to label these kinds of areas “feral” or “ungoverned.” These descriptions fail to account for the informal systems of governance that emerge from the chaos. But since the vast majority of growth in many megacities enters the informal economy and are not accounted for by existing, recognized governments, these alternate control structures are becoming increasingly important. The vacuum left by the neglectful (or incapable) central authority is filled by informal leaders who provide rudimentary services including, water, sanitation, food distribution, means of trade, healthcare and public safety. These leaders often impose taxes on their residents, creating, in effect, a shadow government and society that coexists alongside the formal one. The communities are growing rapidly and, in some places, can emerge to rival the formal government.

If these alternative governance structures are not the target of a military operation, no part of the doctrinal approach to urban operations accounts for them. In fact in many cases military leaders are prohibited from working with these unofficial governors. Author David Kilcullen suggests that most of these alternatives to central authority are "urban street gangs, communitarian or sectarian militias, insurgents, bandits, pirates, armed smugglers or drug traffickers, violent organized criminal networks, vigilantes and armed public defender groups, terrorist organizations, warlord armies and certain paramilitary forces."[35]

Yet by automatically assuming that alternative governance is malignant, leaders are inclined to overlook the usefulness and legitimacy of some of these actors. This view limits thinking about the structure of governance in what is becoming the most likely operating environment for future operations. Not every unofficial governor is evil. In many cases informal leaders are simply filling a gap in governance, and in others they are protecting their constituents' way of life from a central authority that does not agree with them. Consider the Hasidic community in New York City, where city officials rarely enforce the rule of law, instead relying on community leaders to keep the peace.[36] The community has its own justice system, ambulances and robust community councils. They coexist with the surrounding community, and accede to central authority only when necessary. This community and its control structure does not pose a challenge to the government, yet if the military were to apply its doctrinal approach to policing the city, they would ignore this structure and instead insist that the NYPD and others impose "control" over those portions of the city. Or, just as likely, the Army leaders on the ground would communicate almost exclusively with the official government, and only

meet with the Hasidic leaders in order to convince them to subordinate themselves to the central authority.[37]

In other places, competitive control structures may threaten the overall stability of the city. In some of Brazil's favelas, for example, informal leaders rallied communities to resist construction of stadiums for the 2014 World Cup. But in other places, informal governance structures are actually increasing the stability of the city, such as in Mexico City where *autodefensas*—"self-defense militias"—have formed armed opposition to drug cartels.[38] In Lagos, Nigeria, defacto mayors in slums provide for sustenance, education, electricity and an economy in places where the formal government is absent. In 2006, Lagos' informal community leaders formed Lagos Marginalized Communities' Forum (LAMCOFOR), a representative body for over forty-two slum and evicted communities.[39] Anarchy thrives without these emergent governance structures, further threatening urban stability. Successful operations in these environments will require deliberate divergent thinking, yet doctrine limits the commander's opportunities to seek novel solutions.

Since doctrine does not account for alternative control structures, it also does not instruct how to identify who the emergent non-enemy leaders are or where their span of control reaches. Doctrinal approaches tend to oversimplify, labeling people or groups in a binary manner: friend or foe, without regard to the contribution the entity makes to the stability of the overall system.

The military must identify the physical and the social fault lines that divide the city, and who or what has the most influence on them. Current methods of identifying these fault lines can take months or years, relying on person-to-person contact and

trolling social media. Yet the information is readily available in the form of big data, waiting to be analyzed.

Finding Big Data

Humans are now producing more data every year than we previously produced throughout our entire history. In the past two years we have produced a zettabyte of data, which dwarfs the entire record of human civilization.[40] As the world continues to datify[41], vast storehouses of data continue to build up, only a fraction of which is being used. Military and intelligence analysts rarely (if ever) venture into these data sets, often lacking the tools or expertise to properly exploit the data. A significant reason the military doesn't exploit the mountains of data resident in public repositories or with private companies is that it doesn't trust them. Suffering from a form of "not invented here" bias, the military typically does not trust data that it did not generate itself. Yet big data analytics tends to produce insights that are vastly more reliable than traditional methods.[42]

Massachusetts Institute of Technology (MIT) Megacity Logistics Lab studies logistical flow systems in large urban areas.[43] Using data collected by companies through thousands of systems and sensors, Dr. Edgar Blanco identifies "urban channels" that aid companies in streamlining delivery to their final destinations. Companies around the world collect trillions of data points every year on delivery methods and routing, which is then analyzed to search for further efficiencies. This is done in every city in the world, no matter how poor. Companies such as Coca-Cola, PepsiCo, FedEx, and Proctor and Gamble know where every road, trail and sidewalk is, whether a delivery truck can get there, and roughly how much traffic travels on it. Delivery is tracked from distribution to every store, including "nano stores" that service

only 100-150 people per day.[44] Information regarding the routes is updated in near-real-time, allowing leaders to make informed decisions about routing as conditions change with traffic, construction, or other disruptions. The capability to create this kind of knowledge in real time with this level of detail would greatly benefit leaders trying to understand their environment, and would significantly aid tactically planning.

Delivery information is only one type of data that corporations collect and store. The amount of business data that is generated each year is growing exponentially. Even in cities in the developing world companies like Coca-Cola tracks where every drop of potable water comes from, changes in consumption rates, population shifts, economic changes, and many other environmental factors that may affect its operating environment. This data exists now, can be made available, and would change the way landpower leaders see the terrain.

Additionally, maps of local knowledge are more available now than ever before. While military cartographers focus on mapping physical geography, there lies in wait a treasure trove of information that can reveal far more useful information about a location than just its coordinates. As author Stephen Johnson put it, “Anyone can create a map that shows you where streets intersect and where hotels are; we’ve had maps like that for centuries. The maps now appearing are of a different breed altogether: maps of local knowledge created by actual locals. They’re street-smart. They map the intangibles.”[45] Social sites that rate the quality of service or products, or identify unsafe neighborhoods or where potholes go unfilled all provide relevant information about the environment that, when combined with intelligence, can lead to a far richer understanding of the operating environment. This kind of information reveals not only information about the

location itself but about the environmental factors that commanders in urban environments need to be aware of.

Another rich source of data for mapping is collected by one of the biggest online corporations in the world: internet giant Google collects and analyzes well over 20,000 terrabytes of data every day.[46] Google's famous Street-View car has driven on all seven continents, photographing every major (and most minor) trafficable roads in every major city. Additionally, Google is combining street-level and satellite imagery, crowd-sourcing and smartphone geo-data to map buildings in 3D. Local populations provide crowd-sourced inputs to keep up with changes in infrastructure, which has been proven to be highly accurate. These images and models combined with geo-tagged pictures of interiors may soon provide an up-to-date interior and exterior view of every building in the world.[47] Neither civilian nor military intelligence systems can produce anything close to this capability. As one Google official pointed out, "NGA and NSA can't keep up. There are four billion people taking geo-referenced photos and [Google is] going to map all of it."[48] The results thus far have been incredibly accurate, and would prove to be an incredible asset to operators in future urban operations.

Open Source Intelligence centers garner significant amounts of information from posted or published information online, much of it from social media. Exploiting social media requires more than just reading Twitter feeds and following suspected terrorists on their Facebook accounts. From a data perspective, the Tweets themselves carry little value. The true value lies in the metadata, the "information about information" that accompanies each and every Tweet. Each Tweet carries with it a packet of metadata that includes 33 discreet items including user's language, geolocation, and the number

and names of people they follow and those who follow them.[49] This information, when analyzed and mapped, reveals social structures and human behavior elements that were impossible to uncover before.

People are good sources of opinion or perspective, but (as the intelligence manuals point out) they are grossly unreliable. Bias (conscious or unconscious) or self-serving agendas taint the information they provide, whether they are witting providers or not. What people say or write is not always the best source of information as to what is happening in a city, or even what is happening with that individual. What people actually *do* is far more revealing than what they *say*. Where they go, who they connect with, how they move, and where they spend their time and money, are all far more significant indicators of what is really happening. Data Scientist Alex Pentland refers to this kind of data as “digital bread crumbs,” and he uses aggregates of these crumbs to paint very detailed pictures of communities.[50] For example, by aggregating, analyzing and mapping anonymous cell phone data in real time, data scientists are able to leap far beyond mere demographics when describing a city. The analysis reveals subgroups (referred to as tribes) based on behavior patterns. These behavior demographics are far more revealing than the classic demographics based merely on religion or race. These demographics reveal underlying preferences, means of social learning, and drivers of social norms. As Petland describes it, “People within the same behavior demographic have similar food habits, similar clothes, similar financial habits, and similar attitudes towards authority, and as a consequence, they have similar health outcomes and similar career trajectories."[51] This tribal dynamic occurs even if the people within it have never met. Understanding these social tribes offers insight as to how ideas spread and

how large groups of people may respond to new information or action. This understanding can help commanders see the potential responses to operations or adversary action.

Social scientists have often sought to determine how groups of people may act in response to a particular stimulation. However, traditional methods of collecting data about a population relied on sampling. Sampling provided fairly reliable insight about a population in the macro sense, but it was deficient in providing insight into subgroups or at the micro level. With sampling rates typically around n= *a few hundred* (often representing a tiny fraction of the population), the extrapolated information lacked depth. With big data analytics scientists can approach n=*all*, giving them the ability to subcategorize and deeply investigate correlated or anomalous data.

Current intelligence practices do not focus on aggregated big data to reveal social patterns in urban areas.[52] Instead, intelligence practices focus on reducing or eliminating clutter so intelligence personnel can focus collection on just a few targets. Large urban areas make this task extremely difficult, given the incredibly high density of signals. Focusing intelligence efforts typically requires identifying individuals (usually by name) and finding and tracking a specific phone number. This is an incredibly difficult task in cities where mobile telephone subscriptions sometimes outnumber the population. To add to the intelligence community's challenge, there are currently over 100 countries that have more cell phone accounts than actual people, and most of these reside in the major cities.[53]

While useful in tracking individuals, this bias towards micro-analysis, Voice Positive Identification (VPID), and phone numbers associated with specific handsets

doesn't produce the macro understanding necessary to comprehend the city as a system. Further, depending on voice intercepts to determine the zeitgeist is not an efficient or particularly effective way of developing a sophisticated understanding of the urban ecology. Instead, focusing efforts on exploiting the very density that frustrates intelligence collection can provide incredible insight about the social, economic and political networks at play, both formal and informal.

Mobile phones (especially "smart phones") typically reveal an astonishing amount of information about the user, often without their knowledge. One sensing system developed at MIT is designed to continuously collect over 25 phone-based signals, including "location, accelerometry, Bluetooth-based device proximity, communication activities, installed applications, currently running applications, and multimedia and file system information."[54] This information gives the analyst a vast amount of information about the individual, the network, and how the exchange of goods and services take place, as well as potentially mapping the spread of ideas or diseases.

The ubiquity of smart phone technology means that this kind of information will be available to commanders in virtually every operational environment. The markets are already flooded in North and South America, Europe. Much of Asia, and Africa are projected to have over 334 million smartphone connections by 2017.[55] Indeed, the International Telecommunications Union expects 55% of all mobile-broadband subscriptions to be in the developing world in 2015.[56] There are currently almost 7 billion mobile subscriptions worldwide, nearly overtaking the world's population.[57] While access to the data that this brings can be extremely useful in rural areas, it is truly

invaluable in large urban environments where aggregate data can reveal social trends, groupings, and fault lines that give the commander significant clarity about the social and physical landscape. If used correctly, it's like handing the commander a "socio-scope" that allows him to see and track things in real time he could never see before.[58]

Despite the rapid growth of broadband and mobile communications in every city in the world, the number of people communicating on the internet is now dwarfed by the number of things talking on the internet.[59] Machine-to-machine communication has now outpaced human interaction on the web, and carries with it more information about the environment than ever before. Indeed, more than 90% of all the data in the world has been generated since 2011.[60] The proliferation of chips, sensors and communication capability in more and more everyday items has accelerated datification in much of the world around us.

Networked communication is becoming increasingly automated as more automated items join the architecture. The Internet of Things refers to sensors and actuators that are embedded in physical objects that are linked to the internet either through wires or wirelessly.[61] Many of these sensors work largely without human interaction, and produce huge volumes of data. These may include household appliances, automobiles, personal fitness or communication equipment, pacemakers, sensors in public infrastructure, or anything else that connects to the internet. Analyzing the mountains of data generated by these machines can reveal changes in the urban metabolism, social, economic or political changes, and can help predict infrastructure failures that could impact operations. As author Vikto Mayer-Schonberger states in *Big*

Data, "The data can reveal secrets to those with the humility, the willingness, and the tools to listen."[62]

Using Big Data

While the myriad sources of large volumes of data provide ample access to large data sets, they are only useful if they are properly analyzed and the results are used to inform sense-making and decision-making. “Collecting lots of diverse types of data very quickly does not create value,” says an Oracle white paper on big data analytics. “You need analytics to uncover insights.”[63] The purpose of analytics is to gain an objective, deep understanding of important phenomena, and to give leaders fact-based understanding of the environment that goes beyond intuition when making decisions.[64] Properly processing large data sets requires investment in software, hardware and talent.

There are generally three kinds of analytics: descriptive, which reveals information about the past, which includes profiling, segmentation or clustering); predictive, which uses algorithms and models based on historical data to predict the future; and prescriptive, which uses models to suggest courses of action, typically based on an attempt to find optimal solutions. Most businesses applying big data analytics apply all three types, but are increasingly using prescriptive analytics to aid in decision-making.[65] While the Army has begun exploring analytics to better understand and track itself (particularly regarding readiness), there is no indication that any military organization is focused on big data analytics to help them understand cities.

Big data and modern analytics capabilities make it possible to study more complex systems by looking beyond linear relationships and identifying nonlinear relationships among data.[66] Understanding nonlinear relationships can give

commanders and staffs insights as to what the real social undercurrents in cities mean to operations, and what subgroups are driving changes. Data properly analyzed makes it possible to learn which parts of the city generate most of the electronic traffic, which parts initiate electronic or telephonic conversations, and how connected they are to the rest of the city, the rest of the country, or the rest of the world. From this leaders can determine where the true social, economic, religious, or political leaders reside and how much real influence and power they wield. As Pentland says, “Continuous streams of data about human behavior allow us to accurately forecast changes in traffic, electric power use, and even street crime and the spread of the flu."[67] There is no reason these streams of data can’t help the military follow the spread of ideas, particularly extremist ideas that might lead to violent outbreaks.

Analytics can also reveal qualitative assessments regarding the level of integration a city enjoys. For example, a 2011 analytical study on Sao Paulo using only open source data determined that the city has a “highly integrated city center on global level but extremely segregated local centralities.”[68] This dynamic explains the structure of the city, where flow systems and infrastructure radiate from the center where economic activity is centered. The periphery is defined by mono-functional housing, and is separated from other areas by the industry that has emerged along the radials. This means that, with the exception of the city center, most of Sao Paulo is a collection of neighborhoods rather than a highly integrated city. This slows the spread of ideas, norms, and economic gain as flows from a single section rarely spread laterally, instead primarily using the efficient flows systems that benefit the city center. Knowing this may make it possible to temporarily isolate portions of the city without dramatically disturbing

the rest of the population. It also indicates that units operating within an isolated sector of the city will have difficulty getting support from units in adjacent areas as lateral movement is also restricted. These insights emerge from identifying system variables that are highly correlated.

Big data analytics seldom produces information on causation, instead revealing correlations between different data. This makes it descriptive, revealing *what* without the *why*, which can be very useful. Without fixating on causes, decisions become more straightforward: if two things are strongly correlated, does it matter *why* they are? In many cases it does not. Akin to psychologist Dietrich Dorner's "black box" where predictable results are achieved without a full understanding of the relationships that make it so, decisions based on correlation can prove to be the only kind possible.[69] Indeed, in Complex Adaptive Systems (CAS), causation is increasingly difficult to attribute, particularly since so many variables are interdependent. Mapping correlations using big data in this kind of system can help produce predictive models that are extremely useful for decision-making. Further, modeling the complexity of a CAS such as a megacity makes it obvious that applying a reductive hypothesis (such as attempting to find a single "center of gravity") is not only impossible but also undesirable.[70]

Understanding the correlations in a system allows leaders to embark on decision-making without forming a hypothesis about causation. Instead of hypothesis-driven operations (our current model), leaders can apply a data-driven approach. This method may result in dramatically faster decision-making with less bias and higher accuracy.[71]

How the Army might exploit Big Data in Cities

The kinds of operations the Army might be called to conduct in a megacity span the entire Range of Military Operations (ROMO). Regardless of the kind of operation, there are significant advantages to be gained by exploiting big data. In addition to the increased understanding of the urban ecological systems discussed above, big data can provide highly accurate real-time mapping during operations. This has been applied on several occasions during disaster relief operations. For example, Crisis Mappers around the world helped guide rescuers to survivors in the 2010 Haiti earthquake by mapping survivor location based on geo-tagged Twitter and Facebook posts.[72] Data-laden social media feeds also helped drive real-time mapping during 2013 flooding in Germany and during Hurricane Sandy where responders were able to prioritize disaster response using the input.[73] Real-time mapping based on big data may work well for other kinds of military operations as well. Additionally, data sources are not limited to voluntary data provided on social media. Any geo-tagged data contributes to the fidelity of the map, and comes with significantly more information than traditional methods of mapping. These can be collected anonymously, without creating concerns about individual privacy, and can still provide extremely useful data. In the cases for disaster relief, volunteers around the world donated time, expertise and computational capacity to provide relief workers with up-to-date information using anonymous data collected from open sources. To take advantage of this capability, the Army would likely have to replicate this capability itself, particularly for operations other than humanitarian relief. This can be accomplished using home-station units dedicated to providing support to forward-based units, for example. The added benefit for rotational deployments would be dramatically increased situational awareness for subsequent rotations, as follow-on

units that assisted in data analysis would be intimately involved with understanding the environment prior to their deployment.

The Army must improve its ability to gather and analyze big data. Investing in basic and applied research towards higher-level machine learning would unlock the potential of big data to help commanders better understand their environment. New computer systems have made remarkable gains in capability over the past five years. IBM's Watson, for example, uses natural language processing (NLP), automated hypothesis generation, inference and evidence-based hypothesis-scoring to analyze immense data sets almost instantaneously.[74]

Big data analytics can also aid in modeling large urban areas for planning, exercises and decision-making. Currently, operations are planned using digital maps of terrain, with a separate process for investigating socio-economic and cultural factors. With big data-driven modeling, operations can be planned using virtual models that incorporate petabytes of data, providing exceptionally sophisticated feedback to planners. These models have the potential to simultaneously depict the complex interaction between the physical, virtual and social structures of a city. With this capability, these models can be used for exercises, showing leaders how complex the environment is and how their actions might affect the city and its population. Current exercise models for large urban operations use human teams aided by overly-simplistic computer models to adjudicate outcomes of operational moves. This is insufficient for developing leaders who will be successful in future urban environments. Finally, as a decision support tool, advanced urban modeling based on big data can help

commanders understand the impacts of their decisions both as a predictive analytic tool (based on probabilistic outcomes) and as a real-time feedback loop.

A Caution

Big data is essentially descriptive in nature, and often lacks the ability to explain. In other words, it tells you *what* is happening, but without proper analysis it doesn't tell you *why*. While powerful, it has limits, and leaders should avoid being lured into committing the "sin of McNamera" where a leader became so obsessed with the power and promise of data analytics that he failed to appreciate its limitations.[75]

Big data will only improve decision-making if leaders apply it correctly. In reality, few leaders in crisis situations rely heavily on data to inform decision-making. Despite ready access to highly accurate data sets, UN officials in multiple disaster relief operations made decisions based on anecdotes rather than facts. According to a 2013 United Nations Office for the Coordination for Humanitarian Affairs (OCHA) report decision-making was not improved by access to big data analytics.[76] Media reports and political influence still dominated decision-making in most recent crisis responses. This phenomenon happens in military decision-making as well, where leaders tend to rely on experience and intuition despite the availability of data. This is often the result of a lack of trust in either the sources of data or the analytics used to process it. Big data can reveal what is happening, but a leader must still make the decision.

Conclusions and Recommendations

For the foreseeable future, megacities will continue to grow, and enormous amounts of data about them will continue to accumulate. As the Army considers its future operational environments, it must include the megacity environment as a

significant possibility. Studying the modern urban environment using the most advanced analytics and the richest data sets will provide the Army with the best possible understanding of the operational ecology. While still a relatively immature field, big data analytics promises to be incredibly important in shaping the future social, political and economic environment around the world.

The scale and complexity of megacities limits what the Army will be capable of accomplishing within them. It is important for our leaders to have a realistic, well-informed understanding of what the Army can actually accomplish in these environments. While nearly everyone agrees that ground combat operations in a megacity is extremely undesirable, the Army owes it to the nation to find ways of achieving strategic goals in large urban areas. Clearly, current Army doctrine and approaches are inadequate to guide successful operations in increasingly urbanized terrain.

The Army must begin to study megacity environments in earnest. It must expose its leaders to the megacity environment. Developing expertise in urban planning, the science of cities, and big data analytics will accelerate institutional learning. The Army must also invest in research and development that furthers its ability to analyze big data sets and helps it determine which factors are the most relevant in the urban setting. Further, the Army must develop relationships not only with potential partner militaries, but with megacity civil leaders, security forces, and, perhaps most importantly, with their data keepers and data scientists.

The Army must develop big data modeling for simulation training, exercises, and supporting planning and decision-making. Relying on simplistic models reinforces one-

dimensional thinking and reductive hypotheses, and too often amplifies problems rather than resolving them.

Finally, the Army must select, train and develop leaders who think holistically about complex problems in megacities. It must develop leaders who are open to new ideas and are willing to innovate, and who are comfortable operating in uncertain and ambiguous environments. An understanding of big data analytics will help future Army leaders trust the models produced by big data, and ultimately arrive at better decisions within this complex urbanized terrain

Endnotes

[1] Chief of Staff of the Army Strategic Studies Group (CSA-SSG), *Megacities and the United States Army: Preparing for a Complex and Uncertain Future* (Washington DC: US Army, 2014), 1. This is an unpublished report prepared for limited distribution.

[2] Chief of Staff of the Army Strategic Studies Group (CSA-SSG), "A Proposed Framework for Appreciating Megacities: A US Army Perspective." *Small Wars Journal*, April 21, 2014. http://smallwarsjournal.com/jrnl/art/a-proposed-framework-for-appreciating-megacities-a-us-army-perspective-0 (accessed February 22, 2015)

[3] A review of both Joint and Army doctrine reveals a constant call for "systemic thinking" but then only provides reductionist tools for looking at systems. Complex Adaptive Systems (CAS) must be understood both in terms of its parts and as a whole. Emergent properties of CAS are not visible from merely studying the component parts. Army doctrine, in particular, often confuses the term *systemic* (pertaining to or affecting the body as a whole) with *systematic* (giving or using a system or method; methodical). See JP 3-06 *Joint Urban Operations*, JP 2-01.3 *Joint Intelligence Preparation of the Operational Environment*, ATP: 2-01.3 *Intelligence Preparation of the Battlefield/Battlespace*, FM 3-06 *Urban Operations*, ATTP 3-06.11 *Combined Arms Operations in Urban Terrain*, FM 2-91.4 *Intelligence Support to Urban Operations*, and TC 90-1 *Training for Urban Operations*.

[4] Big data refers to data sets typically too large or complex for traditional analysis. For an excellent overview of big data and analytics see IBM's "What is Big Data" at http://www.ibm.com/big-data/us/en/.

[5] Wiki-sourcing is a term used to describe applications that allows collaborative participation in modifying, adding or deleting content. The openness of the content encourages self-correction, often producing highly accurate information. Popular examples include Wikipedia, Linux operating system, and Wikimapia. For further reading see Don Tapscott and Anthony D. Williams, *Wikinomics: How Mass Collaboration Changes Everything* (New York: Penguin, 2006). This is not the same as a "majority is usually right" approach.

[6] U.S. Joint Chiefs of Staff, *Joint Urban Operations*, Joint Publication 3-06 (Washington, DC: U.S. Joint Chiefs of Staff, November 20, 2013), I-6.

[7] JP 3-06, I-7.

[8] JP 3-06, I-8.

[9] JP 3-06, I-8.

[10] Ibid.

[11] Ibid.

[12] United Nations Department of Economic and Social Affairs, "World's Population Increasingly Urban with More than Half Living in Urban Areas," July 10, 2014, http://www.un.org/en/development/desa/news/population/world-urbanization-prospects-2014.html (accessed February 16, 2015).

[13] Ibid.

[14] United Nations Habitat, *State of the Urban Youth Report*, (Nairobi, Kenya: UN-Habitat, 2013), xii.

[15] William Matthews, "Megacity Warfare: Taking Urban Combat to a Whole New Level", *Army Magazine* 65, no.3 (2015): 6.

[16] It is important to note that this report does not call for massive increases in troop strength in order to dominate the environment. This is likely impossible in any event. The author advocates a smarter approach that trades better understanding and better decision making for the mass that traditional doctrines have prescribed.

[17] The author assumes that the approach the Russians used in the 4th Battle of Grozny in 2000 (where fuel-air explosives "along with iron bombs, surface-to-surface missiles with high-explosive warheads, massed artillery and tank fire were used to flatten large sections of the city…") is not a suitable or acceptable option. For more information on the weapons used in Grozny and their effects see Lester Grau, "A Crushing Victory: Fuel-Air Explosives in Grozny 2000", http://fmso.leavenworth.army.mil/documents/fuelair/fuelair.htm.

[18] The ability to understand and exploit the science of cities, both physical and virtual is a dramatic shortfall across the Joint Force. Tactical cyber, non-lethal capabilities, vertical logistics, and the development of urban expert operators are just some of the important areas that warrant serious study. These are beyond the scope of this project.

[19] Ecology is the study of the relations and interactions between organisms and their environment, including other organisms. "Human Terrain" refers to the development of a social science understanding of the population. "Human Terrain Teams" were employed in Afghanistan and Iraq to conduct operationally directed research projects, collect data and develop a common operating picture of the socio-cultural environment. For more information see *US Army TRADOC Human Terrain System*, http://humanterrainsystem.army.mil/index.html.

[20] Samaniego and Moses, “Cities as Organisms: Allometric Scaling or Urban Road Networks.” *Journal of Transport and Land Use*, 1:1 (Summer 2008): 21-39.

[21] CSA-SSG, 31. The medical procedure metaphor is based on an SSG internal working paper produced by Michael Bailey, used with permission.

[22] Adrian Bejan, *Design in Nature: How the Constructal Law Governs Evolution in Biology, Physics, Technology, and Social Organization* (New York: Doubleday, 2012).

[23] In 1996 Bosnia, for example, COL Greg Fontenot’s staff detected massive influxes of external actors in the city of Brcko by monitoring increases in fuel and food consumption. COL (R) Greg Fontenot, interviewed by author, Carlisle, PA, May 15, 2014.

[24] U.S. Joint Chiefs of Staff, *Joint Intelligence Preparation of the Operational Environment*, Joint Publication 2-01.3 (Washington, DC: U.S. Joint Chiefs of Staff, June 16, 2009).

[25] JP 2-01.3, xii.

[26] U.S. Department of the Army, *Intelligence Preparation of the Battlefield/Battlespace*, Army Training Publication 2-01.3, (Washington, DC: U.S. Department of the Army, November 2014).

[27] U.S. Department of the Army, *Plan Requirements and Assess Collection*, Army Training Publication 2-01, (Washington, DC: U.S. Department of the Army, August 2014).

[28] Ibid., 8-4.

[29] U.S. Department of the Army, *Open-Source Intelligence*, Army Training Publication 2-22.9 (FMI 2-22.9) (Washington, DC: U.S. Department of the Army, July 10, 2012).

[30] IBM has launched a series of “Intelligent Operations Centers” in some major cities including Rio de Janeiro, Palava, India and Portland, Oregon. See *IBM Intelligent Operations Center*, http://www-03.ibm.com/software/products/en/intelligent-operations-center/ (accessed March 9, 2015).

[31] Global cities are those considered to be important hubs in the global economic system. See AT Kearney, “2014 Global Cities Index and Emerging Cities Outlook,” 2014, http://www.atkearney.com/documents/10192/4461492/Global+Cities+Present+and+Future-GCI+2014.pdf/3628fd7d-70be-41bf-99d6-4c8eaf984cd5 (accessed February 22, 2015).

[32] See JP 3-06 which states in its opening paragraph regarding planning for urban operations, “Planning for operations that will be conducted in urban environments generally follows the same basic process as planning for operations in other environments.”

[33] In fairness, the doctrine does state that each of the cities is unique and that solutions need to adjust to that uniqueness. Regardless, the intellectual approach in doctrine to understanding the city remains unfazed by uniqueness.

[34] David Kilcullen, *Out of the Mountains: The Coming Age of the Urban Guerilla* (New York: Oxford University Press, 2013), 16-17.

[35] Kilcullen, 126.

[36] New York Police Department official, interview by author, New York, February 10, 2014.

[37] This was the logic applied in Operation Enduring Freedom in Afghanistan where US strategists ignored tribal structures in favor of imposing a Hamiltonian-style federalist system.

[38] Daniel Fisher and Christopher Mercado, "'Competitive Control': How to Evaluate the Threats Posed by 'Ungoverned Spaces'," *Small Wars Journal* (September 17, 2014), 1.

[39] Social and Economic Rights Action Center, "Lagos Marginalized Communities' Forum (LAMCOFOR)," February 17, 2015, http://www.serac.org/programs/community-action-program-cap/community-organizing (accessed February 17, 2015).

[40] Jonathan Shaw, "Why 'Big Data' is a Big Deal," *Harvard Magazine*, March-April 2014, 1.

[41] Viktor Mayer-Schonberger and Kenneth Cukier, *Big Data: a Revolution that will Transform how we Live, Work, and Think* (New York: Harcourt Publishing, 2013), 73. Datification is distinguished from digitization in that data is indexable and thus searchable. Much of what is currently digitized is not indexed but exists as a "dumb" graphic. Automated processes are currently datifying huge volumes of digitized information every day.

[42] Mayer-Schonberger and Cukier, *Big Data*, 4.

[43] For more information on MIT's Megacity Logistics Lab see "Big Data in City Logistics" at http://megacitylab.mit.edu/research/big-logistics-data/.

[44] Dr. Edgar Blanco, interview by author, Cambridge, MA, May 13, 2014.

[45] Stephen Johnson, *Ghost Map: The Story of London's Most Terrifying Epidemic – And how it Changed Science, Cities, and the Modern World* (New York: Riverhead, 2006), 218.

[46] Erick Schonfeld, "Google Processing 20,000 Terabytes a Day, and Growing," Jan 9, 2008, http://techcrunch.com/2008/01/09/google-processing-20000-terabytes-a-day-and-growing/ (accessed February 22, 2015).

[47] Google official, interview by author, Washington, DC, January 29, 2014.

[48] Ibid. NSA refers to the National Security Agency and NGA refers to the National Geospatial-Intelligence Agency, the two US intelligence agencies responsible for cyber and mapping.

[49] Mayer-Schonberger and Cukier, *Big Data,* 93.

[50] Pentland, 8.

[51] Pentland, 141.

[52] This is an unclassified report. The author consulted with multiple intelligence agencies and military intelligence personnel regarding the use of big data analytics and found that none of them are applying it towards understanding cities. It is possible that classified efforts are underway and that the author or the officials interviewed simply weren't aware of them.

[53] International Telecommunications Union, “The World in 2014: ICT Facts and Figures,” April 2014, http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf (accessed February 19, 2015).

[54] Pentland, 222.

[55] Informa UK, “Africa Telecoms Outlook 2014,” 2013, http://files.informatandm.com/uploads/2013/11/Africa_Telecoms_Outlook_Low_resolution.pdf (accessed February 18, 2015).

[56] International Telecommunications Union.

[57] Ibid.

[58] “Socio-scope” is a term used by Alex Pentland in *Social Physics*.

[59] Michael Chui, Markus Loffler, and Roger Roberts, “The Internet of Things,” March 2010, http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_internet_of_things (accessed February 16, 2015).

[60] Science Daily, “Big Data, for Better or Worse.” May 22, 2013, http://www.sciencedaily.com/releases/2013/05/130522085217.htm (accessed February 23, 2015).

[61] Chui, et al., “The Internet of Things”.

[62] Mayer-Schonberger and Cukier, *Big Data,* 5.

[63] Oracle, “Big Data Analytics: Advanced Analytics in Oracle Database,” March 2013, http://www.oracle.com/technetwork/database/options/advanced-analytics/bigdataanalyticswpoaa-1930891.pdf (accessed February 20, 2015).

[64] Thomas H. Davenport, “Analytics 3.0,” *Harvard Business Review*, December 2013, https://hbr.org/2013/12/analytics-30/ (accessed February 19, 2015).

[65] Ibid.

[66] Mayer-Schonberger and Cukier, *Big Data,* 61.

[67] Alex Pentland, *Social Physics: How Good Ideas Spread – the Lessons From a New Science* (New York: Penguin Press, 2014), 143.

[68] Claudiu Forgaci, “Tools for Urban Analysis: Analysis of Sao Paulo’s Spatial Configuration and its Implications Defining Polycentricity,” *EMU Technology*, 2011, 17. http://issuu.com/claudiuforgaci/docs/120223_technology_report_cforgaci_rev2 (accessed February 20, 2015).

[69] Dietrich Dorner, *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations* (Cambridge, MA: Perseus Books, 1996).

[70] Ibid., 89-92.

[71] Mayer-Schonberger and Cukier, *Big Data,* 55.

[72] Meier, Patrick. “How Crisis Mapping Saved Lives in Haiti.” *National Geographic*, July 12, 2012.

[73] Laura Gurski, “Using Big Data in Disaster Relief,” *World Economic Forum*, January 13, 2015, http://www.atkearney.com/web/davos-2015/home/-/blogs/using-big-data-in-disaster-relief (accessed February 23, 2015).

[74] *IBM Watson*, http://www.ibm.com/smarterplanet/us/en/ibmwatson/what-is-watson.html (accessed February 23, 2015).

[75] Mayer-Schonberger and Cukier, *Big Data*, 169.

[76] UN Office for the Coordination of Humanitarian Affairs, “Humanitarianism in the Network Age,” March 6, 2013, http://www.unocha.org/node/11528 .