warontherocks.comhttp://warontherocks.com/2016/01/bringing-big-data-to-operations-in-mega-cities/

Bringing Big Data to War in Mega-Cities

Robert Dixon

As the U.S. Army prepares for the future, it has become increasingly aware that operations are more and more likely to take place in large cities. The number and size of cities continues to grow, and they are quickly becoming the dominant form of human habitation. Belligerent actors, aware of the West's growing anxieties about collateral damage, have good reason to place forces in or around cities. Further, advanced sensing and weapons systems employed by modern militaries make hiding in remote areas of the world less and less attractive to non-state enemies of advanced powers.

America's enemies see the advantages of the seemingly impenetrable clutter that dominates the modern city. The Army's current approach to learning about this environment is to seek the diamonds scattered amidst this clutter. What we are missing, though, is that the clutter itself is the jewel. Enormous amounts of readily available data can reveal more about a city, its

population, and the nefarious actors residing there than we could have imagined before. To truly understand this environment the Army must fundamentally change its approach to understanding the environment: It must adopt a holistic approach enabled by big data analytics.

The Army, however, seems hesitant to embrace 21st-century data analysis, instead relying largely on the same micro-level methods it has used for decades. This must change if the Army wishes to maintain the ability to "see first" and "understand first" in the modern urban arena.

The Urban Challenge

Political leaders and security forces have always gathered data to better understand their environment. Yet cities have presented a particular challenge to data gathering with their constantly changing infrastructure, myriad subcultures and ample places to "hide in plain sight." Censuses and geographic mapping are centuries-old techniques, and have always been time-consuming and lacking in accuracy. Time delays between gathering and analyzing data and presenting conclusions have too often produced unreliable and out-of-date information, making well-informed, real-time decision-making difficult at best. Even in today's operations, the U.S. Army still relies heavily on traditional methods of individual (scout, leader observation, etc.) as well as platform (imagery and intelligence) observation, two-dimensional mapping, and population surveying. In the past, these methods were deemed sufficient as there were no alternatives.

But the modern urban environment is changing, further challenging past methods of seeking understanding. Rapid urbanization across the globe has given rise to megacities (defined as cities with more than 10 million residents) and mega-regions in which major cities "grow together" forming regions of dense population that stretch hundreds of kilometers and can encompass over 100 million people. The rapid growth in urban areas produces more demand on the infrastructure and flow systems, more waste, and increased urban density. It also increases the likelihood that the Army will be tasked to operate there.

While the term "megacity" has a specific definition, there is nothing magical about that specific distinction. Some urban areas with fewer than 10 million people pose as significant a challenge to operating forces, while other, larger urban areas may offer more straightforward approaches. The scale, density, connectedness, complexity, and threat all contribute to the nature of the problem and the potential solutions. Size is but one of the relevant factors.

Traditional methods of collecting data about a population largely rely on sampling, often by surveying individuals within that population. Sampling provides fairly reliable insight about a population in the macro sense, but is deficient in providing insight into subgroups or the micro level. With sampling rates typically around n=*a few hundred* (often representing a tiny fraction of the population), the extrapolated information lacks depth. With big data analytics, analysts can now approach n=*all*, giving them the ability to subcategorize and deeply investigate correlated or anomalous data. More importantly, data derived from involuntary sources (e.g. cell phone data or financial transactions) reveal a more honest picture than survey-driven sampling: It shows what people *actually do* as opposed to what they *say* they do.

Behind the Big Data Curve

Technological advances over the past decade have changed one aspect of the modern city more than anything else: Cities are producing enormous quantities of data and data analysts are learning how to use it in new ways. Recent progress in big data analytics, the proliferation of automated sensors, the ubiquity of mobile technology, the democratization of information, and the "datification" of nearly every social, economic, and logistical transaction provide previously unimaginable insight into modern urban ecology. Big data analytics may have made many traditional methods of collection and analysis all but obsolete.

Cities and corporations produce much of their own data as a byproduct of normal operations, and have invested heavily in acquiring relevant data produced by others. Capturing and databasing enormous data sets, as well as data sharing among stakeholders, are becoming easier and more commonplace with technological advances. Indeed, there are thousands of publicly available data sets that are available to any user wishing to access them. Failure to develop capabilities to exploit this data practically ensures that the Army will be far behind other actors in understanding the environment.

One of the most significant sources of data lies in mobile communications. The ubiquity of smartphone technology means that enormous amounts of data are generated continuously, even in the least developed cities. There are currently almost 7 billion mobile subscriptions worldwide, nearly overtaking the world's population. While access to the data that this brings can be 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 leaders significant clarity about the social and physical landscape. If used correctly, it's like handing the commander what MIT's Sandy Pentland calls a "socio-scope" that allows him to see and track things in real time that he could never see before.

Governments and major corporations routinely collect billions of data every day — even within the poorest cities — and the quality and volume is increasing exponentially. Humans are now producing more data every year than we produced throughout our entire history. Urban leaders use big data analytics to plan infrastructure and improve service delivery. Corporations use big data to increase value by making information transparent and more accurate, and as a tool to understand and segment value-chain stakeholders. Both government and commercial users of big data aim to achieve the same goal: better understanding of their environment. Sadly, the Army currently lacks the resources, expertise, approaches, or seemingly even the desire to investigate and exploit the reservoir of information available in modern cities. This must change.

As the world continues to datify, vast storehouses of data become vaster. While military and intelligence analysts sometimes venture into these data sets, they are typically searching for individual nodes or linkages, attempting to find the virtual needle in the big data haystack. What they are ignoring is the value of understanding the dynamics of the haystack itself. This "micro-bias" dramatically limits the value inherent in large data sets. While big data analytics tends to produce insights that are vastly more reliable than traditional methods, the Army seems stuck in the "way we have always done it."

Plugging In, Switching On

As the scale of modern urban areas continues to increase and the [absolute number of land forces in Western armies continues to decline](#), militaries must come to terms with their limited abilities to operate in urban environments. Current Western expectations about the conduct of war require modern militaries to seek ways of accomplishing their goals without resorting to targeting populations and the infrastructure that supports them. A robust, sophisticated understanding of the urban ecology is necessary to identify the appropriate pressure points to apply force against. This understanding, along with modeling and operational feedback loops, can provide future commanders with a learning mechanism that not only helps identify targets, but also suggests the right tools (lethal or nonlethal) for servicing them. Moreover, sophisticated models maintained by real-time sensors and big data analytics enable a self-awareness commanders have long needed: near instantaneous feedback on the effects of their operations.

Properly equipped, a commander can now gather, analyze, map and model a city's infrastructure, population dynamics, and sub-group behavioral patterns in a matter of days or weeks instead of months or years. More importantly, once gathered, modeled, and monitored, commanders can observe changes in these systems in real time as data streams are updated continuously. Most importantly, this could be accomplished with a minimal military presence in the city itself.

While powerful, big data has limits and leaders should avoid being lured into committing the ["sin of McNamara"](#) where a leader becomes so obsessed with the power and promise of data analytics that he fails to appreciate its limitations. Big data will only improve decision-making if leaders apply it correctly. Data analytics typically reveals *correlation*, but does not speak definitively about *causation*. Leaders still have to decide how best to use the information garnered from analytics.

In reality, though, over-reliance on data analytics is rarely a problem. In fact, [few leaders in crisis situations rely heavily on data to inform decision-making](#). Media reports and politics still dominate decision-making in most crisis responses. Leaders tend to rely on experience and intuition to make decisions, despite the availability of data.

Moving Forward

The Army must study megacity environments in earnest. It must expose leaders to the megacity environment as often as possible. 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.

The Army must move beyond crude models and 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. While the Army has conducted some simulation experiments in and around megacities in the past few years (most notably the [Army's annual Unified Quest](#) experiment), these efforts have been far too simplistic, often aggregating large numbers of

disparate social groups into a few manageable ones, and wishing away many of the complexities inherent to the modern urban environment.

Certainly the Army has begun working towards developing better understanding and approaches to megacities. The work done at TRADOC's Army Capabilities Integration Center (ARCIC) Future Warfare Division and CSA's Strategic Studies Group, among others, have done much to highlight the Army's challenges in these areas. Army thinkers have begun debating the importance and potential effects urbanization brings to military operations, evidenced by the Spring 2015 *Parameters* issue featuring three articles on the topic. Outside the Army, USSOCOM's work with Caerus in developing structured approaches that attempt to account for the relationship between physical, socio-economic, and operational aspects in urban areas shows some promise. Indeed, Caerus' insistence that trying to use Excel or PowerPoint to analyze urban systems is wrong-headed is a huge step in the right direction. Yet, sadly, there is little evidence that the Army is interested in developing the capability to conduct rigorous big data-driven analysis, instead relying largely on the same reductionist models (see PMESII) that limit holistic thinking. This must change.

Finally, the Army must select, train and develop leaders who think holistically about complex problems in large urban areas. It must develop leaders who are open to new ideas, willing to innovate, and 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 terrain.

Colonel Robert Dixon is an Army Engineer officer currently serving as the Corps Engineer for I Corps. He has served as a strategist and planner at the Combined Joint Task Force, Division, and Combatant Command levels, and served on the Chief of Staff of the Army's Strategic Studies Group. Colonel Dixon is a graduate of Florida Institute of Technology, American Public University, the School of Advanced Military Studies, and the Army War College where he was a member of the Carlisle Scholars Program. The views expressed here are his own and do not necessarily represent those of the U.S. Army or Department of Defense.