Big Data for Good:

Insights from Emerging Markets

Rajesh Chandy

London Business School, London, UK

rchandy@london.edu

Magda Hassan

University of Warwick, Warwick Business School, Coventry, UK

magda.hassan@wbs.ac.uk

Prokriti Mukherji

King’s College London, School of Management & Business, London, UK

prokriti.mukherji@kcl.ac.uk

Big Data for Good:

Insights from Emerging Markets

Abstract

This paper examines how innovations involving big data are helping solve some of the greatest challenges facing the world today. Focusing primarily on the developing world, we explore how the large volumes of digital information, increasingly available in these contexts, can help decision makers better understand and better address problems as big as poverty, illness, conflict, migration, corruption, natural disasters, climate change, and pollution, among other areas. We argue that the information vacuum that still exists in many developing countries makes the potential for impact from big data much greater in these contexts, and we outline what practitioners and academics can do to make a difference with big data.

Keywords: big data, innovation, leapfrogging, data poverty, emerging markets, social impact.

Big Data for Good:
Insights from Emerging Markets

1. Introduction

Innovations have the greatest impact when they solve intense needs for large numbers of people. This truism prompts at least three fundamental questions as we consider the nature, scope, and likely impact of innovation in data-rich environments. How can the opportunities offered by big data be effectively harnessed to create positive impact? Where are the greatest opportunities for impact from innovations that harness big data? Who are the primary actors who can contribute to and benefit from innovation involving big data?

As the volume, variety, and velocity of data in the world has exploded, a growing scholarly and practitioner literature has sought to address these questions (cf. Einav and Levin 2014; Laney 2001; Marr 2016; McAfee and Brynjolfsson 2012; Wedel and Kannan 2016). Many discussions of big data, implicitly or otherwise, focus on opportunities for impact in advanced economies, and in contexts that are, in relative terms, among the most advantaged in the world. A logic underlying this focus is the same as that purportedly offered by Willie Sutton when asked why he robbed banks: "because that's where the money is" (Sutton and Linn 1976). These contexts are seemingly awash in new data: most of the data being generated today is being generated in developed world contexts, where the digital sources of such data are ubiquitous. A great deal of attention has focused on the challenges of storing, analyzing, and creating commercial value out of this new data (see Varian 2014; Wedel and Kannan 2016). Indeed, the commercial opportunities for profiting from innovations that harness such data are also possibly the highest in the most advanced economies, since customers in these contexts have the greatest wherewithal to pay for these innovations.

This article offers a different perspective on the how, where, and who questions mentioned in the opening paragraph. Rather than focus primarily on the opportunities offered by big data innovation to yield commercial impact, this article focuses on the opportunities for big data innovation to have a social impact. Rather than focus primarily on developed world contexts, we focus on some of the least advantaged parts of the world. We highlight the potential for big data to do good in emerging markets. Rather than focus primarily on commercial actors, we note the possibilities for impact through the complementary activities of government entities, NGOs, academic researchers – as well as commercial entities. In doing so, we seek to make three contributions.

First, we emphasize that the goals toward which big data innovation can be channelled are not merely commercial. Big data can also be harnessed to solve some of the most profoundly important problems facing humankind today (cf. Desouza and Smith 2014). We illustrate the possibilities for big data to do good by noting the impact of such data in addressing problems such as: 1) forced migration, 2) disease, 3) poverty and economic stagnation, 4) ethnic divisions, and 5) ecological and environmental crises.

Second, we highlight the value of examining the role of big data in developing countries (cf. Burgess and Steenkamp 2006; Sheth 2011; Sudhir 2016). Because these contexts have historically suffered from data poverty rather than an abundance of data, and because many basic human aspirations remain unfulfilled in these contexts, the intensity of need – for the largest numbers of people – is possibly the greatest in these contexts. We describe the challenge of data poverty, and note how big data innovation has the potential to transform lives in these markets. Furthermore, we propose a new 3Vs framework – comprised of Validation, Visualization, and Verification – that could form a research agenda for scholars who seek to make an impact through big data innovation.

Third, we demonstrate the importance of the full chain of activities involved in creating value from big data (see Figure 1), and go beyond the focus in the literature on storing, analyzing, and creating commercial value. In part because the storage, analysis and profit possibilities of big data have been covered elsewhere (McAfee and Brynjolfsson 2012; Varian 2014; Wedel and Kannan 2016), we do not address these activities in detail here. Instead, we emphasize the links highlighted in green in Figure 1: we point to the importance of innovation in identifying, integrating, disseminating, and applying new sources of data to execute actions that in turn generate product, service, process, and business model innovations – and ultimately, to be impactful. As we note in the section below, in contexts marked by data poverty, the creation, integration, and the execution of actions based on rich data cannot be taken for granted.

2. Data Poverty

The data context in many developing countries has historically been the polar opposite of big data (cf. Laney 2001): limited volume, limited variety, and limited velocity. Policy and practice often relies on government statistics, which are generally sparse and outdated. Even when data on crucial variables is available, it is often inaccessible: storage may be poor, sharing might be restricted, and incompatibility in formats may make pooling across different sources difficult. Rather than being digitized, data is still stored in stacks of paper in dusty spaces that at any point could be subject to damage by fire, floods or dirt. Additionally, institutions that hold these data are sometimes secretive, and not willing to share it with outside researchers (or indeed, sometimes with rival departments within their own governments). Linkages and exchanges between institutions that hold different types of data that might supplement each other are lacking, thus preventing a full picture of reality from emerging.

Even when adequate data is available and accessible, it is frequently of poor quality. Often, the data relies on surveys that are predicated on proper comprehension on the part of the respondent (Grosh and Glewwe 2000). But survey comprehension among poorly educated respondents may be poor. Surveys are often conducted in situations that are ripe for response bias. A solitary respondent may be asked to respond to sensitive questions (about income, health, gender issues, etc.) while surrounded by groups of their family members or neighbors.

Finding skilled surveyors and local research staff can be an additional challenge, with predictable consequences for data quality. Surveyors may not have proper training, and the responses they record may be inaccurate and biased. Or they may be paid very little and not properly supervised; the temptation to make up numbers is non-trivial. Even if the enumerators are properly trained and conscientious, respondents might find them quite intimidating due to the difference in education and social status between them. Moreover, there may be institutional pressures (from political bosses, for example), to systematically over- or under-report information. Indeed, the reality of data contexts in emerging markets may be summarized by the old Woody Allen joke from *Annie Hall* about a conversation between two elderly women. One of them says: "Boy, the food at this place is really terrible." The other responds: "Yeah, I know; and such small portions."

Yet, as we noted earlier, people's needs in emerging markets are intense – and often very basic. These contexts are the global hotspots for some of the most urgent problems facing humankind: poverty, disease, migration, pollution, and climate change, among others. So the potential for innovation - through new approaches, new insights, and new actions - is huge. Indeed, as we outline in the section below, emerging markets are sometimes leapfrogging developed markets in the creation of, and the impact of, big data innovation.

This paper proceeds as follows. We first illustrate, through a series of case studies, how big data can be used to address big social and environmental challenges in developing countries. We present research questions that could not have been addressed in the absence of dramatic recent increases in data volume, variety, and velocity. We then extrapolate from these questions, and discuss the nature of the technological changes that now allow decision makers in developing countries to leapfrog from data poverty to big data, and permit innovative solutions to the aforementioned challenges. We end with a discussion of the actions that researchers and practitioners should take to further capitalize on the opportunities for doing good that are offered only because of the big data revolution.

3. Big Data for Good: Some Illustrations

Question 1: Where do people migrate after natural disasters? In January 2010 Haiti, and particularly the town of Porte au Prince (PaP) was hit by one the worst earthquakes in its history. This disaster left thousands of people dead and even more without any shelter, food or water. As a result more than 500,000 of PaP's population left the city during the days following the earthquake, hoping to find shelter in other parts of the country. The National Civil Protection Agency collected data on migration using traditional data collection methods: asking people for the place they intended to go to. This data quite accurately captured the number of people who left PaP, but was extremely inaccurate in predicting where people actually went since the intended destinations were quite different from the actual destinations.

Big data made accessible by Digicel, one of Haiti's biggest mobile operators (which covered about 90% of PaP's population at the time) made it possible to accurately track migrants' location using SIM card login data from their network. The data, which included calling records of 2.8 million SIM cards and 282 million registered calling locations, was offered to academic researchers by Digicel. The blue dots on Figure 2 represent the number of migrants received by different communities all over the country. Figure 3 shows the discrepancy between the mobile operator data estimate and the National Civil Protection Agency estimate of the number of migrants received per community. Consider the implications of this discrepancy. Inaccurate estimation in this case could have led to wasted (scarce) emergency resources in the form of food, health, and shelter in some communities and unavailability of such resources in communities that desperately needed them (Bengtsson et al. 2011). This example illustrates the value of data charity by a commercial entity to academic researchers who then integrated this data with that from the Haitian government and the United Nations Population Fund. Innovation in the form of mobile telephony - a relatively new technology in the Haitian context that created a rich new source of data - and new ways of integrating, analyzing, and representing the data offered policy makers, charities, and other disaster relief agencies actionable insights that would have been impossible to obtain otherwise.

Question 2: Can big data innovation predict epidemic outbreaks? A couple of months after the Haiti earthquake, specifically in October 2010, cholera broke out in the town of St. Marc. Those fleeing their homes in the town could have spread the disease to farther flung areas, making it difficult to target interventions to prevent its spread. Once again, SIM card data offered a means to track the movement of St. Marc's population and to identify potential new outbreaks of the epidemic. Thus, Big data helped in identifying those communities that were at higher risk of cholera outbreaks due to such migration, thereby allowing preventative and curative medical interventions to prevent further spread of the disease and contamination of other communities (Bengtsson et al. 2011).

Question 3: Can big data assess economic growth? When targeting some of the poorest countries in the world, both business planners seeking new markets, and policy makers making decisions regarding these contexts face difficulties in accurately measuring economic growth. Even the poor and unreliable data that does exist is often only available at high levels of geographic aggregation: for example, at the national level. Rarely is it available in disaggregate form, in part because this data is based on nationwide surveys of representative populations, and too few data points exist in any given locality to make reliable inferences about that locality. Henderson, Storeygard, and Weil (2009) pioneered an ingenious method that allows inferences at the local level. These authors make use of night-time light intensity data as a proxy for economic growth. Specifically, they use data that is collected by US Air Force weather satellites that observe every location on the planet every night at some point between 8:30 and 10 P.M. local time, and is made publicly available by the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center (NGDC).

Figure 4 shows the light intensity on the Korean peninsula between 1992 and 2008. The image clearly shows the rapid economic development that South Korea has seen during this period. In contrast, little improvement has happened in North Korea, which has stagnated economically. It is particularly worth noting that in the absence of this big data, getting an accurate image of the economic performance of North Korea would have been almost impossible, given the secretive nature of the regime there.

While it is important to get an accurate measure of macro-economic performance indicators such as GDP, it is equally important to develop a more granular micro-economic figure of the development of nations and markets. In this context, Michalopoulos and Papaioannou (2012) show that light density data can also be used as a proxy for micro-economic indicators such as per capita income, piped water, level of education and sewage system at the local (few square kilometers) level in Africa. Vast amounts of data, originally collected in large part for military intelligence purposes, now supplemented by commercial entities applying innovation in the form of low cost satellites and specialized high resolution cameras, and integrated and analyzed by academic, non-profit, and commercial entities, offer the potential to transform our knowledge of economic and social development.

Question 4: What is the impact of pre-colonial ethnic institutions on economic development? The light density data described earlier allowed Michalopoulos and Papaioannou (2013) to answer a fundamental (and difficult to research) topic: the role pre-colonial ethnic institutions on economic development in Africa today. These authors first map the pre-colonial boundaries of ethnic homelands in Africa. They then code the complexity of pre-colonial political institutions from 0 (petty chiefdoms) to 4 (large complex states). The authors show that areas that were part of complex states in the pre-colonial era are 3 times more economically developed (greater light intensity) today than areas that were petty chiefdoms in the pre-colonial era. Once again, big data created by innovations in satellite and imaging technology, and applied using new mapping and analysis techniques, provides fresh insights on the long reach of the past on the present and future.

Question 5: How many people live below the poverty line? Estimates of how many people live below the poverty line vary widely. Yet, targets for poverty alleviation such as the UN's revised Millennium Development Goals, as well as those by aid agencies and local governments all require, at a minimum, reasonably accurate assessments of the state of the problem. Data in the places worst affected by poverty are too often inaccurate, inaccessible, or at far too high a level of aggregation. In contexts such as these, big data on mobile phone usage might offer a solution. Blumenstock, Cadamuro, and On (2015) make use of data on the entire population of Rwandan mobile phone users to infer poverty levels. They validate the use of mobile phone data as a measure of individual-level poverty in three steps.

First, the authors conducted 860 phone surveys from which they developed a composite wealth measure. They then matched the phone survey responses with the mobile phone data; the mobile phone data correlate strongly with a composite wealth index developed using the phone survey data. The second step was then to match the mobile phone data with a government census of 1200 households. This procedure again indicates that mobile phone data offer a good proxy for household wealth. Finally, they construct a high-resolution map of Rwanda using call records of billions of interactions among 1.5 million subscribers. The map shows the predicted level of wealth in a given cell of 5-15 subscribers living in that region. This analysis goes far beyond existing sources of data on poverty in its accessibility and reach, and allows much more targeted actions from those seeking to help the poor.

Question 6: Can big data help address illegal deforestation and global warming? Global deforestation accounts for one-fifth of the world's annual emissions and greenhouse gases, an amount that is much greater than the greenhouse gas emissions produced by the global transport sector. Given the sensitivity of this issue – and the political power and money available to some of the worst offenders – official and self-reported data on this issue is unreliable. As an alternative approach, Burgess et al. (2011) study deforestation via satellite images obtained from NASA of Indonesian forests from 2001 to 2008. By combining Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery (which captures deforestation at a 250-meter by 250-meter resolution annually for the entire country) with Geographic Information Systems (GIS) data on district boundaries, they show that the more administrative fragments each Indonesian district is divided into, the greater the extent of illegal deforestation. Greater levels of competition for illegal payoffs between officials in rival administrative units lead to a race to the bottom, and the more fragmented the district in question the greater the competition for such payoffs. The images of Indonesia over time (Figure 5) that are made possible by big data show, in poignant and dramatic detail, what no traditional government inquiry or NGO investigation could possibly show: the extent of permanent destruction to the earth's ecosystem.

Question 7: Can big data help businesses do good – and make profits?

In the above discussion, we have largely focused on possibilities for policy makers and researchers to serve as agents of change through big data innovation, with commercial entities largely playing the role of passive but generous sources of data. But we would be remiss if we did not point out that the opportunities for socially impactful innovation via big data are of course not restricted to such non-commercial entities. A single example may offer a glimpse of the possibilities for big data to make a difference among the poor – while making a profit.

M-Kopa is a for-profit company that offers solar energy based, mobile money enabled products and services to the roughly 65% of East Africans who do not have access to grid electricity. Lacking electrification, most Kenyans, Tanzanians, and Ugandans spend roughly 50 US cents (or roughly 15-20% of monthly income for the average household) on kerosene for lighting. Kerosene is a fairly poor light source, and kerosene lamps are a major cause of indoor air pollution and accidental fires. M-Kopa's solar kits, which offer clean solar lighting for multiple rooms (as well as a radio and mobile phone charging facility) represent an ingenious example of big data being applied to business model innovation (Demil et al. 2015) and sustainable innovation (Varadarajan 2015).

Because many East Africans cannot afford the up-front cost of their solar systems, M-Kopa offers customers a pricing scheme that is a variation on a traditional installment payment plan. Customers can acquire the M-Kopa system for $50. Once installed, they need no longer spend the 50 US cents they would otherwise (on average) spend on kerosene; instead, they apply this amount to payment for the M-Kopa system. Each 50 US cent payment offers 24 hours of solar energy on a pay as you use basis. All payments are made via mobile money (which is close to ubiquitous in Kenya and Tanzania today). An embedded SIM card monitors the usage and status of each solar system in use, and communicates in real time with M-Kopa's cloud-based databases. After 365 daily payments, the customer owns the solar system outright. At this point, use of the system is free and unlimited for the life of the system. Moreover, M-Kopa makes use of its knowledge of each customer's payment and usage history (which it has thanks to machine-to-machine communications) to offer credit as well as upgrades to higher-powered solar systems and to consumer durables such as televisions to credit worthy customers. Customers with high need and greater ability to pay (as judged by their observed usage and transaction history) are those targeted for these new products and services offered by the company.

As of December 2016, M-Kopa had already sold over 400,000 solar devices across four countries, and had access to data – through machine-to-machine communications between each of these devices and their data servers - on hundreds of millions of mobile payments per year, 1 million battery readings per day, as well as sunshine readings (for potential weather forecasting applications) at hundreds of thousands of (GPS mapped) locations. The M-Kopa solution demonstrates the possibility to profitably change data poverty in financial records, weather records, and energy consumption records.. M-Kopa executives, who are based in the UK and in Africa, generate so much data in Africa that they are currently the largest users of Microsoft's cloud-based data storage systems in Africa. They are now implementing additional new products and services (beyond the higher powered lamps, televisions, and loans that they already offer) that are designed to capitalize on the unique data that the company owns on their customers needs and ability to pay. Big data enables M-Kopa to transform their customers' access to clean, cheap energy, new information sources, and credit availability. More generally, big data offers new possibilities for customers who would otherwise likely be unreachable by companies that do not have this data, and new possibilities for business model innovation by these companies (Bharadwaj and Noble 2016; Demil et al. 2015).

4. Discussion: Big Data for Good

In this paper, we posit that some of the greatest opportunities for impact through innovation in big data exist in contexts that have historically been plagued by data poverty. By addressing social and development goals in emerging markets, we highlight – through case studies of academic research - the value of innovations that generate big data, and of innovations that harness big data. We conclude by discussing the forces that enable these opportunities for innovation, and ways in which academics and practitioners can make effective use of these forces.

Leapfrogging data: Technological innovation is arguably the primary force that allows those in the poorer regions of the world to benefit from big data. The story of how those in developing countries leapfrogged over fixed line telephones to embrace mobile telephony is well known (Chandy and Ramdas 2013). Indeed, many of the big data opportunities described in this article are the fruits of this revolution. But satellites, biometric tools, machine-to-machine communications, cheap sensors, and cloud-based storage are all now creating possibilities that were non-existent before. Indeed, it is conceivable that poorer regions of the world could leapfrog from small data (traditional methods of data generation) to big data (high volume, velocity, and variety of data) without going through the same process of evolution that happened in wealthier regions that have many more legacy data sources (cf. Sheth 2011; also see Hauser, Tellis, and Griffin 2006; John, Weiss, and Dutta 1999).

What features of these technologies might enable this data revolution? First, these technologies enable easy connectivity: information is easy to transfer, and therefore to collect and analyze. Second, they are often characterized by continuity: be it be mobile masts, satellites, Global Positioning Systems, or cloud-based storage, these technologies are (almost) always switched on. This feature not only leads to high volume and variety of data, but also much greater granularity of detail in data. Third, these technologies offer the benefit of incidentality: in many cases, these technologies were never really intended for the purposes for which they are now being used. Who would have thought, when mobile phones first came to the market, that the traces they leave could help prevent the spread of epidemics? Or that satellites that had their origins in military surveillance could help reduce the destruction of rain forests? This incidentality often makes them cost-effective sources of data, since the investments behind them are cross-subsidized.by others (e.g., the military, or consumers in developed countries).

Implications for academics. Academic researchers have an important role to play in helping the world harness the potential for big data innovations such as those we have described here. In the paragraphs below, we propose a new 3Vs framework that could form part of a research agenda for future researchers seeking to improve the veracity of big data: Validation, Visualization, and Verification. We then emphasize the wealth of opportunities that innovations in data-rich environments offer for researchers to make a real difference in the lives of many.

1. Validation: The task of demonstrating the validity of the various sources of big data that are now becoming available is a crucial one. Much of this data relies on proxies of the constructs of interest to policy makers and practitioners (e.g., light intensity or mobile phone usage as a proxy for poverty or development). How well do these proxies represent the real world phenomena of interest? A satisfactory answer to this question is necessary before major actions or investments are made based on such data. Academics can play a central role in developing the answer, not only through empirical validation, but also by developing and applying theories and frameworks that offer testable hypotheses and guide empirical analysis (Burgess and Steenkamp 2006; Sheth 2011; Yadav 2010).

2. Visualization: Better techniques to visualize big data will increase their value as well as their accessibility. However, many academic researchers – especially those who study poorer regions of the world – came of age during a data-poor era. Their training often focuses on qualitative research, or on statistics involving relatively small samples. The next generation of scholars (as well those from previous generations who seek to remain on the cutting edge of research) might find that training in data visualization can pay off greatly in creating impact. Indeed, all of the studies discussed in this paper rely on visualization to make their

respective points. For example, the pioneering work of Michalopoulos and Papaioannou (2013) was published in *Econometrica*, a journal that typically features a high ratio of Greek letters to English letters. But this paper barely features any mathematical equations, relying almost entirely on visual representations of large volumes of a large variety of data to offer unique and compelling insights.

The richness of data available today opens new vistas for data representation. Now is a good time to revisit the exhortations of scholars who have urged researchers to apply creative ways to visualize data (Holbrook 1997; Ozimec, Natter, and Reutterer 2010; Tufte 1983). Asterisks that show significance levels of model coefficients are fairly useless when sample sizes number in the millions. Visual representations can help simplify and more effectively communicate insights from large volumes of rich data, thus making it more likely that these insights will have impact (see Sawyer, Laran, and Xu 2008).

3. Verification: Spurious correlation is a big risk when conducting analyses with big data. Rigorous tests of causal inference are therefore especially crucial when utilizing such data. The use of scientific methods to infer causality is an important contribution that academic researchers can make for the effective use of big data (Rubin 2008). Research designs that incorporate exogenous shocks or randomized controlled trials could be invaluable to this effort, and these are often easier to incorporate in developing world contexts that are subject to many exogenous shocks, or where randomized interventions are often not cost-prohibitive (e.g., Blumenstock, Fafchamps, and Eagle 2011; Anderson, Chandy, and Zia 2017).

More broadly, big data represents a golden opportunity for innovation researchers to bring the social impact of innovation to the forefront of research on the topic (see Kolk and Lefant 2015; Mick 2007; Nakata and Weidner 2011; Webster and Lusch 2013; Wilkie and Moore 1999). Academic researchers can contribute new knowledge and insight at each link of the value chain outlined in Figure 1. Indeed, most of the case studies of academic research discussed in this article go beyond the links in red in Figure 1: their contribution is in the identification and integration of big data, in addition to storage and analysis. Researchers can also help practitioners understand how to better disseminate - within and across organizations – the novel insights generated from big data. Moreover, innovation researchers can help those seeking to do good to more effectively execute innovations based on the troves of big data now being created in formerly data-poor contexts. By doing so, innovation researchers can contribute solutions to the pressing challenges that face the world.

Implications for practitioners and policy makers. Practitioners and policy makers can build on the momentum that is now beginning to form behind big data in places where decisions previously had to be made based on little information and a large dose of faith. First, we suggest that the push to build infrastructure in places that did not have much before should encompass not only physical infrastructure, but also data infrastructure. Such infrastructure should encompass all the elements of the value chain involved in applying big data for impact. For example, the trillions of data points being created as part of universal identification projects such as the Aadhaar project in India and through mobile money services in East Africa are offering a glimpse into the lives of the poor in a manner that is unprecedented in scale and scope (see Economides and Jeziorski 2017; Muralidharan, Niehaus, and Sukhtankar 2014). Many more such projects are possible and desirable. Government entities, profit seeking firms, and NGOs can all contribute to building the various elements of the value chain in these contexts and beyond.

Second, practitioners and policy makers could ensure that data does not live in isolation. Part of the power of big data is through the insights that can be generated from bringing a large variety of different data sources together. For such data integration to happen, linkages via hubs (for example, facilitated through Application Programming Interfaces) and common standards are crucial. Creating such linkages while preserving data security and privacy requires concerted and thoughtful effort; such efforts will require engagement between governments, businesses, and academics alike. Business can also actively engage in data charity, through which they contribute internal data to others seeking to fulfil social objectives (as for example in the mobile phone cases noted earlier in the paper).

Finally, no data, no matter how big it is, will have an impact unless practitioners and policy makers are responsive to it. Impact therefore requires investment in data generation in emerging markets, in storage, integration, analysis and dissemination capabilities within and across markets and organizations, and a culture that promotes data-enabled innovation in products, services, processes, platforms, and business models (Marr 2016; McAffee and Brynjolfsson 2012). It requires leadership that has a commitment to data driven innovation, as well as the motivation to do better by doing good (cf. Bharadwaj and Noble 2016). Often, this requirement implies a need for first hand experience in two domains: data driven decision-making, and social impact. The former domain is likely to sensitize leaders to the value of investments in data capabilities, and the latter is likely to sensitize them to the shared value that can be created at the intersection of business, government, and society in contexts where many fundamental needs remain under-fulfilled (Kramer and Porter 2011; Wedel and Kannan 2016). Given the relative novelty of these domains, and the opportunities for impact that are inherent in them, those practitioners, policy makers, and academics who create and successfully harness these opportunities will not only serve as active agents of change in their own contexts, but will also serve as sources of inspiration to others far beyond.

5. References

Anderson, Stephen, Rajesh Chandy, and Bilal Zia, 2017. "Pathways to Profits: The Impact of Marketing versus Financial Skills on Business Performance". Working Paper.

Bengtsson, Linus, Xin Lu, Anna Thorson, Richard Garfield, and Johan Von Schreeb. 2011. 'Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti'. *PLoS Med* 8 (8): e1001083.

Bharadwaj, Neeraj, and Charles Noble. 2016 "Innovation in Data-Rich Environments," *Journal of Product Innovation Management*, 32 (3), 476-478.

Blumenstock, Joshua, Gabriel Cadamuro, and Robert On. 2015. 'Predicting Poverty and Wealth from Mobile Phone Metadata'. *Science* 350 (6264): 1073–1076.

Blumenstock, Joshua E., Nathan Eagle, and Marcel Fafchamps. 2016. 'Airtime Transfers and Mobile Communications: Evidence in the Aftermath of Natural Disasters'. *Journal of Development Economics* 120 (May): 157–81.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. 'The Political Economy of Deforestation in the Tropics'. *Quarterly Journal of Economics* 127(4): 1707-1754.

Burgess, Steven and Jan-Benedict E.M. Steenkamp. 2006, Marketing renaissance: How research in emerging markets advances marketing science and practice, *International Journal of Research in Marketing,* 23, pp. 337-356.

Chandy, Rajesh, and Kamalini Ramdas. 2013. 'The M in Future'. *Business Strategy Review* 24 (1): 30–43.

Economides, Nicholas, and Przemyslaw Jeziorski. 2017. 'Mobile Money in Tanzania'. *Marketing Science* (forthcoming).

Einav, L. and Levin, J., 2014. Economics in the age of big data. *Science*, *346*(6210), p.715-721

Goldenberg, J., and Oreg, S., 2007. Laggards in disguise: Resistance to adopt and the leapfrogging effect. *Technological Forecasting and Social Change*, 74(8), pp 1272−1281.

Grosh, Margaret, and Paul Glewwe. 2000. 'Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study, Volume 3'. *World Bank Publications*. https://ideas.repec.org/b/wbk/wbpubs/15195.html.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. 'Measuring Economic Growth from Outer Space'. *American Economic Review* 102(2): 994–1028.

Holbrook, M.B., 1997. Stereographic visual displays and the three-dimensional communication of findings in marketing research. *Journal of Marketing Research*, pp.526-536.

Kolk, Ans and François Lenfant. 2015. Cross-Sector Collaboration, Institutional Gaps, and Fragility: The Role of Social Innovation Partnerships in a Conflict-Affected Region. *Journal of Public Policy & Marketing* 34:2, 287-303.

Laney, Doug. 2001. '3D Data Management: Controlling Data Volume, Velocity and Variety'. *META Group Research Note* 6: 70.

Marr, B., 2016. Big Data in practice: How 45 successful companies used big data analytics to deliver extraordinary results, New York: Wiley.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. 'National Institutions and Subnational Development in Africa'. *Quarterly Journal of Economics* Vol 129:1 p 151-213.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. ‘Pre-Colonial Ethnic Institutions and Contemporary African Development’. *Econometrica* 81 (1): 113–152.

Mick, D.G., 2007. The end (s) of marketing and the neglect of moral responsibility by the American Marketing Association. *Journal of Public Policy & Marketing*, *26*(2), pp.289-292.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2014. “Building State Capacity: Evidence from Biometric Smartcards in India.” NBER Working Paper 19999.

Nakata, Cheryl, and Kelly Weidner, 2011. Enhancing New Product Adoption at the Base of the Pyramid: A Contextualized Model. *Journal of Product Innovation Management*, 29 (1), 21-32.

Ozimec, A.M., Natter, M. and Reutterer, T., 2010. Geographical information systems-based marketing decisions: Effects of alternative visualizations on decision quality. *Journal of Marketing*, *74*(6), pp.94-110.

Rubin, D.B., 2008. For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, pp.808-840.

Sheth, Jagdish, 2011., "Impact of Emerging Markets on Marketing: Rethinking Existing Perspectives and Practices, *Journal of Marketing*, 75 (July), 165-82

Sudhir, K. 2016. “Editorial - The Exploration-Exploitation Tradeoff and Efficiency in Knowledge Production.” *Marketing Science* 35(1): 1-9.

Sutton, Willie and Edward Linn.1976. *Where the Money Was*. Viking Press, New York.

Tufte, E.R. and Graves-Morris, P.R., 1983. *The visual display of quantitative information* (Vol. 2, No. 9). Cheshire, CT: Graphics press.

Varadarajan, Rajan. 2015. ‘Innovating for Sustainability: A Framework for Sustainable Innovations and a Model of Sustainable Innovations Orientation’. *Journal of the Academy of Marketing Science*, 1–23.

Varian, H.R., 2014. Big data: New tricks for econometrics. *The Journal of Economic Perspectives*, *28*(2), pp.3-27.

Wedel, Michel, and P. K. Kannan. 2016. ‘Marketing Analytics for Data-Rich Environments’. *Journal of Marketing* Vol. 80, No. 6, pp. 97-121.

Webster Jr, F.E. and Lusch, R.F., 2013. Elevating marketing: marketing is dead! Long live marketing!. *Journal of the Academy of Marketing Science*, *41*(4), pp.389-399.

Wilkie, W.L. and Moore, E.S., 1999. Marketing's contributions to society. *The Journal of Marketing*, pp.198-218.

Yadav, Manjit S. 2010, “The Decline of Conceptual Articles and Implications for Knowledge Development,” *Journal of Marketing*, 74 (January), 1–19.

Figure 1: Impact from Big Data: A Value Chain

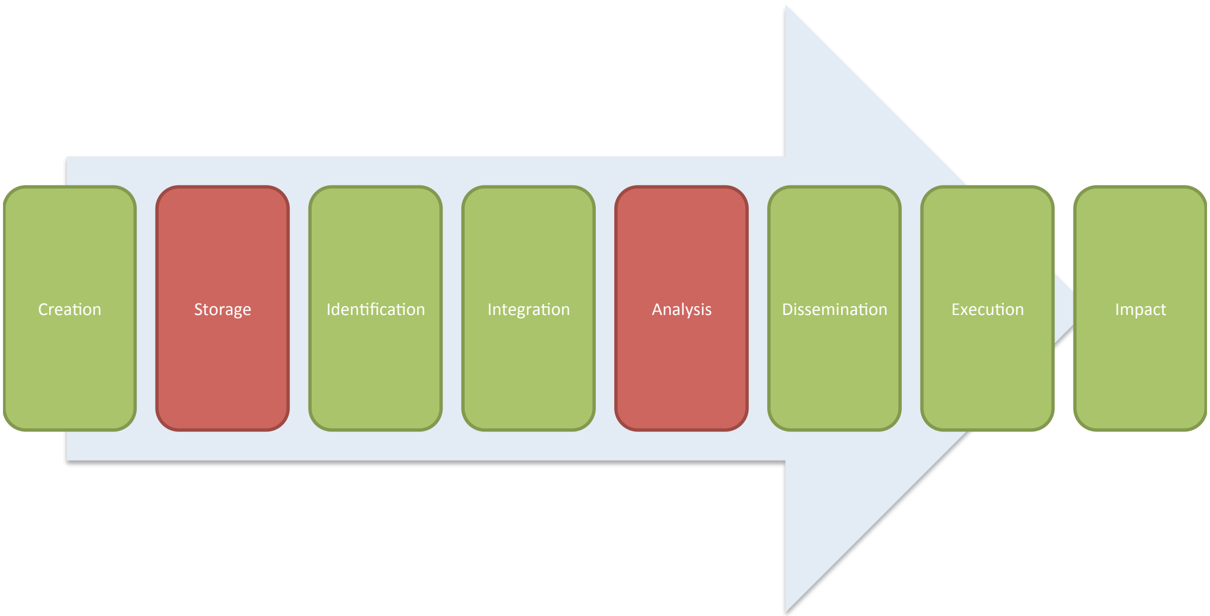
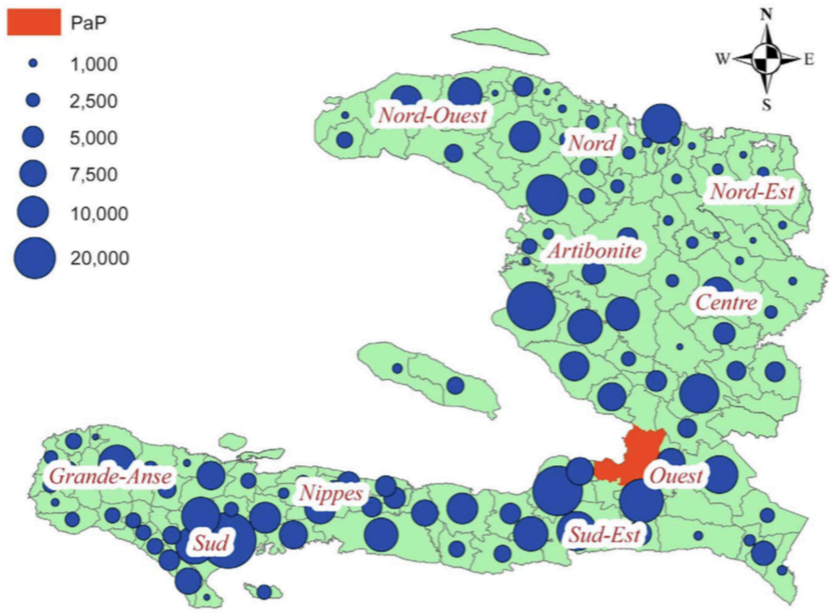
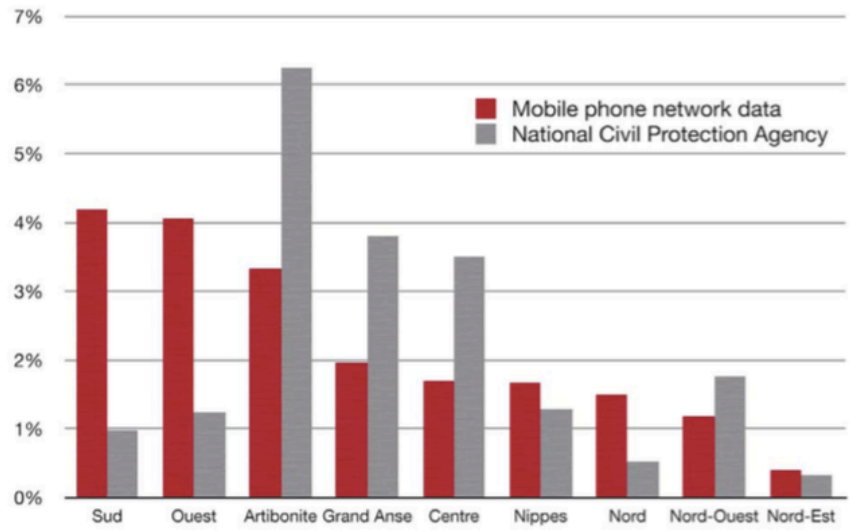


Figure 2: Post earthquake distribution of Port au Prince population (Jan 31, 2010)



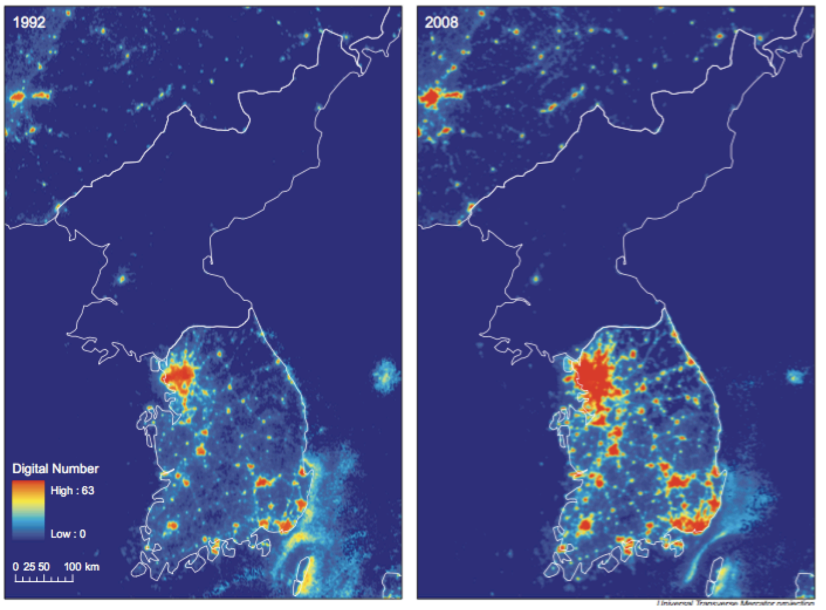
Source: Bengtsson, L., Lu, X., Thorson, A., Garfield, R., & Von Schreeb, J. (2011). Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti, PLoS Med 8(8).Marr

Figure 3: Discrepancies in estimates of the proportion of Port au Prince persons who had left Port au Prince by Febuary 17, 2010, by destination province



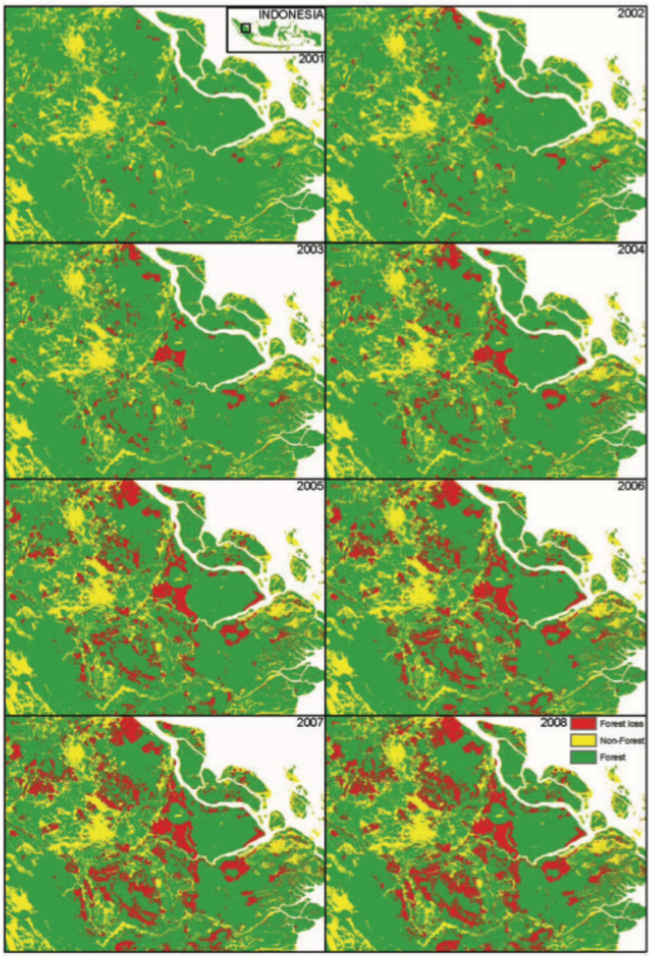
Source: Bengtsson, L., Lu, X., Thorson, A., Garfield, R., & Von Schreeb, J. (2011). Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti

Figure 4: Economic Growth as Viewed from Outer Space: Korea (1992-2008)



Source: Henderson, J. V., Storeygard, A., & Weil, D. N. (2009). *Measuring economic growth from outer space*. National Bureau of Economic Research.

Figure 5: Degree of illegal deforestation in Indonesia (in red): 2001 – 2008



Source: Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. (2012). 'The Political Economy of Deforestation in the Tropics'. *Quarterly Journal of Economics* , 1707–1754