I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

More Than Data: Working with Big Data for Civics

SARAH WILLIAMS*

According to the International Data Corporation, there were 2.8 zettabytes of data produced in 2012, a number expected to double every year until 2020.[1] A zettabyte is roughly equal to 250 billion DVDs or 36 million years of HD video.[2] This unprecedented growth of data has generated excitement for its ability to reshape the way we live. However big data in its raw form cannot perform on its own; rather it is how data is transformed and operationalized that can change the way we see the world. More specifically, data can be used for civic action and policy change by communicating with the data clearly and responsibly to expose hidden patterns and ideologies to audiences inside and outside the policy arena. Communicating with data in this way requires the ability to ask the right questions, find or collect the appropriate data, analyze and interpret that data, and visualize the results in a way that can be understood by broad audiences. Combining these methods transforms data from a simple point on a map to a narrative that has meaning. Data is not often processed in this way because data analysts are not familiar with the techniques that can be used to tell stories with the data. This essay details strategies for operationalizing data to expose issues and generate policy debates.

* Assistant Professor Department of Urban Studies and Planning, MIT, and Director Civic Data Design Lab, MIT.

[1] International Data Corporation Website, accessed June 27, 2014, https://www.idc.com/.

[2] Ibid.

I. DATA AND CIVICS

Recent excitement around big data's potential to change the way we see the world has created an enthusiasm for applying it to civic action and policy change. In 2014, The White House conducted a three-month study on the topic. It reported that big data has the potential to change the way we make decisions not only in government, but also in all sectors of the economy.[3] The White House report follows major investments by mainstream tech companies. IBM developed its "Smart Cities" and "Smarter Planet" advertising campaign in 2008 to promote the use of technology and data to analyze the problems of cities.[4] In 2010, IBM created a partnership with Rio de Janeiro to make it a model "Smart City."[5] During that same year, Cisco launched its "Smart and Connected Communities" program, which is based on using data analysis and web-based interface programs to connect cities through technology.[6] Also in 2010, The Economist magazine developed a series of stories called the "Data Deluge," which explained the possibilities of using data to develop strategies for almost anything, including cities. The hype around these projects generated a fascination in the popular press and media for what the analysis of big data can offer cities.[7] More recently, Microsoft launched its CityNext program, which is meant to draw from its world-wide network of technology experts to make cities better places.[8] In the last few years leveraging big data has become big business and tech companies want a piece of the action.

---

[3] John Podesta, "Findings of the Big Data and Privacy Working Group Review." *The White House Blog*, May 1, 2014, http://www.whitehouse.gov/blog/2014/05/01/findings-big-data-and-privacy-working-group-review.

[4] IBM Smart Cities Web Site, accessed June 27, 2015, http://www.ibm.com/smarterplanet/us/en/smarter_cities/overview/.

[5] Natasha Singer, "Mission Control, Built for Cities: I.B.M. Takes 'Smarter Cities' Concept to Rio de Janeiro," *New York Times*, March 3, 2012.

[6] Pete Swabey, "IBM, Cisco and the Business of Smart Cities: How Two of the IT Industry's Largest Companies Plan to Rewire Urban Living," *Information Age*, February 23, 2012, http://www.information-age.com/industry/hardware/2087993/ibm-cisco-and-the-business-of-smart-cities.

[7] "The Data Deluge," *The Economist*, February 27, 2010.

[8] "Microsoft CityNext: Coming to a City Near You," Microsoft News Center, accessed July 31, 2014, http://news.microsoft.com/2013/07/10/microsoft-citynext-coming-to-a-city-near-you/.

The term "Big Data" is relatively new, but interest in using data to analyze and determine city policies is not.[9] Shannon Mattern, a cultural theorist, traces curiosity about using data for urban planning to modernist notions of "cities-as-machine" and our fascination with the development of cities as efficient systems.[10] Data analysis was also central to technocratic planning, which was developed during early 1910s and 20s with the 1917 Manhattan zoning ordinance being one of its first success stories. Later post-war planners of the 50s and 60s used computer models, largely created for transportation engineers, to make decisions. These planners depended on data and models to generate planning policies, and according to Peter Hall, "Anything that could not be expressed in numbers was inherently suspect."[11] Data was central to these movements because it gave decisions a sense of legitimacy as the use of data was often equated with truth.[12] Little consideration was given to the fact that the data interpretations were heavily biased, by those generating and collecting, the information.[13]

The failures of modernist ideals in cities led to a backlash against using data as the sole basis to make decisions for civic change. Jane Jacobs, an influential civic activist in the late 60s, along with other civic actors, including Herbert Gans, led this charge.[14] Jacobs believed human-centered narratives were essential for understanding the economic and social needs of cities.[15] She fought using data-driven

---

[9] Shannon Mattern, "Methodolatry and the Art of Measure: The New Wave of Urban Data Science," *Design Observer: Places*, accessed July 31, 2014, https://placesjournal.org /article/ methodolatry-and-the-art-of-measure/; Anthony M. Townsend, *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia* (WW Norton & Company, 2013); Sarah Williams, "The Responsive City : The City of the Future Re-Imagined from the Bottom-Up" in *Emergent Urbanism*, ed. Tigran Hass (England: Ashgate Publishers, 2014).

[10] Shannon Mattern, "Methodolatry and the Art of Measure: The New Wave of Urban Data Science," *Design Observer: Places*, November, 2013, https://placesjournal.org/ article/methodolatry-and-the-art-of-measure/.

[11] Hall, Peter, "The Turbulent Eighth Decade: Challenges to American City Planning." *Journal of the American Planning Association* 55, no. 3 (1989): 275-282.

[12] John Forester, *Planning in the Face of Power* (Oakland: Univ. of California Press, 1988).

[13] Rob Kitchin, "The Real-Time City? Big Data and Smart Urbanism," *GeoJournal* 79, no. 1 (2014): 1-14.

[14] Peter Hall, "The Turbulent Eighth Decade: Challenges to American City Planning," *Journal of the American Planning Association* 55, no. 3 (1989): 275-282.

[15] Roberta Brandes Gratz, *The Battle for Gotham: New York in the Shadow of Robert Moses and Jane Jacobs* (New York: Nation Books, 2011).

policies as evidence for building highways because the results often severed or completely removed whole neighborhoods in cities.[16] Jacobs believed that the removal of these communities lacked an acknowledgement of the importance of people and their networks in the larger economy of the city.[17] She and other activists wanted the voice of the public to be included into the decision making process.

Inaccurate models became the key failure of data driven planning policies in the 60s. The construction of these models was flawed because of the lack appropriate data to develop them, the use of biased information, and not testing whether the data results fit people's experiences.[18] A great example of this is the 1970s era model developed by the RAND Corporation for the New York City Fire Department, to determine the efficiency of its fire house network. Basing the model almost solely on response time data, the analysis lead to the recommendation of removing several fire stations in the South Bronx. The data model lacked consideration of additional data factors such as, gridlock, politics, multiple simultaneous fires (and therefore the need for back-up units), and the socio-economic status of the neighborhoods considered for closure.[19] The result caused an overload on the fire stations that remained in the Bronx, and it is estimated that half a million people were displaced by fires during this time period. The prevalence of fires was caught on video during the 1977 World Series game and caused ABC's commentator Howard Cosell's famous remark: "There it is, Ladies and Gentlemen, the Bronx is Burning."[20] The phrase was later used as the title of a book authored by Jonathan Mahler on politics and life in 1977 New York City. [21] The RAND model did not cause the rampant fires in the Bronx during the 70s but it did incorrectly assess how the closures would affect the other firehouses in the network, thus creating an overloaded system.

---

[16] Ibid.

[17] Janes Jacobs, *The Death and Life of Great American Cities* (New York: Random House LLC, 1961).

[18] Anthony M. Townsend, *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia* (New York: WW Norton & Company, 2013).

[19] Ibid.

[20] Jonathan Mahler, *Ladies and Gentlemen, the Bronx is Burning: 1977, Baseball, Politics, and the Battle for the Soul of a City* (New York: Macmillan, 2006).

[21] Ibid.

The misguided models of 1950s and 60s have largely been forgotten with the excitement around harnessing the power of big data for cities. Much of the agenda is being set by large corporations such as IBM, Cisco, and Microsoft which can make a profit by applying their skills to big data analytics in cities. Yet these analysts, often not trained in urban planning or a related field, are not familiar with how data models have been used inappropriately in the past. Given the caveats for using data in planning, particularly the complexity of telling balanced stories, it is essential that anyone working with data learn from the past to generate more nuanced models that can be used for change in the future. My research generates projects that work with data for civic action and policy change. Through investigating these projects, I have come to recognize specific strategies for working with data to create action. These include visualization, human-centered data collection strategies, sharing data, incorporating both quantitative and qualitative data analysis, acquiring private data to tell stories, and ultimately building diverse teams to tell narratives with data. Ultimately these strategies can help guide the future of data analytics in the era of big data.

II. VISUALIZING DATA

Data visualizations are great tools for advocating for policy change because they hold the legitimacy of data, express something that is hard to see, and convey complex ideas to the public. Numeric data stored in databases, .txt files, and spreadsheets makes it hard for the average person to read and interpret meanings behind numbers. Graphic representations of data synthesize this information to make it legible to diverse audiences. Data visualizations must at once be highly selective, presenting one or two ideas, while also engaging a reader in the complexity of the problem being addressed. The selective nature of data visualizations is what makes them effective tools, but also represents the biggest caveat to their use. This is because data visualizations are often used as unbiased evidence. However, their production includes selective decisions by the data collector and designer, which can create biased stories.[22] Data visualizations must be read through the eyes of those who generated the data and developed the representations to truly understand their objectives. Ultimately, visualizations have the power to capture our attention and start a dialog about civic issues.

---

[22] Rob Kitchin, "The Real-Time City? Big Data and Smart Urbanism," *GeoJournal* 79, no. 1 (2014): 1-14.

Million Dollar Blocks is an example of a project that uses data visualizations to expose the policy issues behind America's criminal justice system. Using governmental records the project developed maps that expose city blocks where over a million dollars was spent to incarcerate the people who lived there. The project shows how we spend federal dollars to incarcerate people from concentrated neighborhoods; yet those same neighborhoods have limited social services that would help addresses the systemic issues which cause high incarceration rates.[23] The images captivate because they help bring prison policy to the scale of the city block–something that everyone can understand. The data visualizations allow the public to read the evidence in the maps and decipher their own interpretations. Ultimately the maps were used as evidence to support congressional funding for prisoner re-entry programs, showing that visualizations were easy to use as evidence to support policy.

The Million Dollar Blocks message was at once simple and complicated. It clearly showed the mass amounts of money being spent on incarceration; yet the data had been simplified to tell a story. The data itself, through its collection process, revealed the combined impact inflicted unknowingly through the spending decisions made pursuant to diverse sets of intertwined policies. These decisions were exposed through the visuals. The maps were graphically alarming, so much so, they have been included into the permanent collection of the Museum of Modern Art (MoMA), in New York City where many audiences have experienced the message they present: We spend Millions of Dollars to incarcerate people in America, and the cycle of people in an out of prison has become a big business. Million Dollar Blocks helped bring this policy issue to a larger public and generate a debate around the topic.

---

[23] Laura Kurgan, *Close Up at a Distance: Mapping, Technology, and Politics* (Cambridge, MA: MIT Press, 2013).

Image 1: The above image of the Million Dollar blocks project was used to help argue for prisoner re-entry funding in congress. Image provided by the Spatial Information Design Lab.

III. COLLECTING DATA

"Do-it-Yourself" data collection projects can generate data where formal institutions have failed to do so, empower the public by giving them ownership over the data collection process, and teach data

literacy through the act of collection.[24] Adapting mobile technology for data collection provides alternatives to the limitations of more traditional data. Almost everyone owns a mobile phone, and these little computers have the ability to gather data about nearly anything in real time. Projects that utilize this technology are often framed as "citizen sensing" or "participatory sensing" and have been discussed not just for their method of base data collection but also for their potential to measure and monitor environmental conditions, such as air quality.[25] Collecting data outside formal institutions can help put new issues on the map, identify otherwise unmeasured or uncollected data, challenge the criteria by which existing data has been collected, and correct misleading or inadequate information.

Citizen sensing projects often bring civic issues to the attention of the larger public and start a debate. One example is a project developed by the Spatial Information Design Lab and the Associated Press to expose air quality issues in China by measuring it during the Beijing Olympics. Air quality sensors were attached to the Associated Press reporters' cell phones and provided real-time data on the air quality conditions in Beijing. The results showed levels of particulate matter, the fine particles in the air that cause asthma and other respiratory problems, which are often significantly higher than the World Health Organization standard in developing countries.[26] The sensor data provided the only air quality figures available during the Olympics. The Chinese government's "Air Quality Index" was uninterpretable because no information was provided on how the index was developed. Reading the index meant relying on government interpretation of a "good" or "bad" air quality.[27] Beijing Air Tracks, as the reporters' project came to be known, showed how DIY technologies can be used to provide data where the government is not.[28] The lessons learned from this experiment go beyond exposing

---

[24] Eric Paulos, R.J. Honicky, and Ben Hooker, "Citizen Science: Enabling Participatory Urbanism," in *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*. ed. Marcus Forth (Hershey, PA: Information Science Reference, IGI Global, 2009).

[25] Jeffery A. Burke, et al., "Participatory Sensing," *Center for Embedded Network Sensing* (2006).

[26] Sarah Williams, "Beijing Air Tracks: Tracking Data for Good" in *Accountability Technologies Tools for Asking Hard Questions*, ed. Dietmar Offenhuber and Katja Schechtner (Vienna, Austria: AMBRA, 2013).

[27] Ibid.

[28] Ibid.

air quality conditions in China as the devices developed showed that ubiquitous computing, such as mobile phones, allow us to take control of information and use it to advocate for change.

Image 2: Photos taken of Associated Press reporters collecting air quality data. Image provided by Sarah Williams.

IV. OPENING DATA

Governments, largely in North American and Europe, have begun the experiment of sharing their data. The result has been an explosion of civic-based applications, the forging of new partnerships between civic organizations, and an increased involvement in civics by the technology community.[29] Open data allows governments to generate city infrastructure outside formal governmental structures, creating new possibilities for innovation.[30] Overall, open data has increased

[29] B. Goldstein and L. Dyson, eds. *Beyond Transparency: Open Data and the Future of Civic Innovation* (San Francisco: Code for America Press, 2013).

[30] F. Rojas, *Transit Transparency: Effective Disclosure through Open Data* (Cambridge, MA: Transparency Policy Project, Harvard Kennedy School, 2012).

information sharing, which in turn helps to generate new partnerships, new innovations, and civic action.

While many countries are pushing for open data, it is still tightly controlled by governments across the globe, particularly in developing countries, where access to information represents power. Developing data outside these formal government institutions and providing it openly can foster new relationships between the government and those interested in working towards policy change.[31] Countries in the Global South have the potential to have increased benefit from open data; yet data, if it exists, is often locked away within research institutes, NGOs, and government organizations. [32] Leveraging this hidden resource may be just what these countries need in order to make informed decisions about their urban development. However, the culture of information sharing has not been developed in many countries in the Global South or within NGO's themselves which are producing much of the data.

Nairobi, Kenya is an example of a rapidly developing city where data is not often freely shared. Since 2006, I have teamed up with colleagues to investigate the causes and possible solutions to Nairobi's debilitating congestion problems. In the process of our work, my team has generated data and offered it up freely for anyone to use. The results have generated increased trust between our partners in Nairobi, new relationships with NGO's and the Government, and engagement of Nairobi's robust civic technology community, which has built applications on top of the data.[33]

V. MIXING QUALITATIVE OBSERVATIONS WITH ANALYTICS

Data is not solely quantitative in nature. It also incorporates human narratives, images, video, and observations or qualitative data. Responsible arguments about civic topics include both qualitative and quantitative data analysis. This mixed approach is important as it

---

[31] Sarah Williams, Elizabeth Marcello, and Jacqueline M. Klopp, "Toward Open Source Kenya: Creating and Sharing a GIS Database of Nairobi," *Annals of the Association of American Geographers* 104, no. 1 (2014): 114-130

[32] Rajiv Thakur and Sharma Madhuri, "GIS and Challenges to Planning and Development Applications in Peripheral Regions," in *Planning and Socioeconomic Applications*, eds. J.D. Gatrell and R.R. Jensen (Netherlands: Springer, 2009) 125-137.

[33] Sarah Williams, Elizabeth Marcello, and Jacqueline M. Klopp. "Toward Open Source Kenya: Creating and Sharing a GIS Database of Nairobi," *Annals of the Association of American Geographers* 104, no. 1 (2014): 114-130.

helps to verify quantitative data analysis. Perhaps more importantly, it helps to identify the interests of the people as Jane Jacobs advocated in the 60s. In other words, rather than looking at results on a map from afar, data analysts need to go to the places their quantitative data highlights and hear the stories from those who are living in the situation under study. Interviews, images, and observations can expose missing variables in analytical models and help highlight ways to make those models more accurate. There are always multiples vantage points to any policy investigation, and to hear all those voices it is necessary to look at more than numeric data analysis by combining interviews and images.

City Digits, a project developed to teach data literacy to high school students by allowing them to collect, explore and form opinions about a civic topic, shows the power of bringing together qualitative and quantitative data. A web-mapping tool, developed for the curriculum, and piloted in Bushwick's School for Social Justice in New York City, allows students to explore maps that show data about who wins and loses in the lottery. The data highlights certain communities that spend a greater proportion of their income on lottery. It also looks at the social demographic characteristics of communities where lottery stores generate the most revenue. The tool allows students to collect interviews from lottery players and retailers in the field and add that data immediately to the map. After the students work through the curriculum, they are asked to synthesize the data they collected and use it to develop an argument about whether they think the lottery is good or bad. The resulting stories are added to the web tool, allowing anyone to see the student's interpretations of the data. The process of developing the stories allows the students to bring together multiple viewpoints on how the lottery affects their community and helps them form their own opinion on a civic topic.[34] The process teaches them how multiple forms of data can be used to form an argument about a specific idea and generate a debate around a social issue in their community.

---

[34] Sarah Williams, Erica Deahl, Laurie Rubel, Vivian Lim, "City Digits: Developing Socially-Grounded Data Literacy Using Digital Tools," *Journal of Digital Media Literacy*, December 15, 2014.

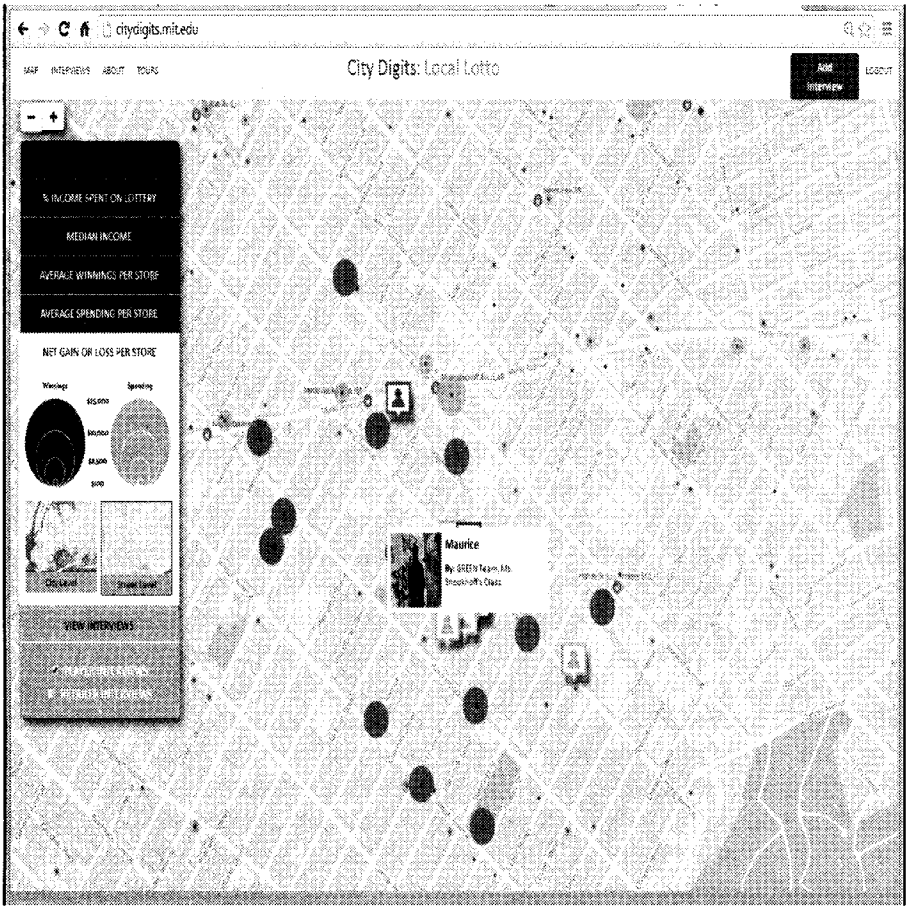

Image 3: City Digits: Local Lotto web tool allows youth to compare quantitative and qualitative data to form opinions on a civic topic. The above screen shot shows a map where students see how much people win or lose by bodega and interviews with lottery players they took in the field. Image provided by Sarah Williams.

VI. PRIVATE DATA USE FOR PUBLIC DISCLOSURE

Cities have traditionally been represented through the filter of governmental data, yet social media and sensors affixed to almost everything offer new perspectives of city life. What is not widely understood is that the digital data we sometimes share unknowingly is often easy to access and can be downloaded, analyzed, and applied to understanding the places we live. Social media is particularly interesting as it collects the impressions of city inhabitants through their own geo-registered sentiments. Public opinion pages, online real estate sites, and social media websites generate gigabytes of data at a

much finer level of detail than traditional governmental sources. Many of these sites ask users to categorize the places they visit by giving them a way to specify what types of business they are interacting with.[35] This data can then be mined from social media sites, which raises many privacy concerns, but at the same time holds much potential. For example, in countries where government data is hard to obtain or does not exist, using crowd-sourced data can allow for the collection of data on land use and real estate prices. Likewise, it is clear that data originally generated from private companies can be used to describe and analyze the city.

Analyzing digital data traces from private companies, such as social media, can create positive change. The "We Are Here Now" Project which uses social media data to describe six global cities, Beijing, Mumbai, Rio, Mexico City, New York, and Moscow, highlights the possibilities of using social media to describe cities. It was groundbreaking because it was one of the first studies to demonstrate the ability to capture geo-registered social media data for whole cities by creatively using the sites API's to tap their vast databases. The study showed that social media traces can tell us about the economic dynamics of city life. Spatial analysis performed on multiple cities highlighted economically driven land use patterns. Moscow showed a prevalence of Mega-Malls and Tokyo's, a focus on transit-oriented development.[36] Understanding urban patterns using the data was particularly helpful in Mexico City were my research team had been trying to access land use data but could not obtain the files. The foursquare business data collected in Mexico, while not comprehensive because it is crowd- sourced, acted as a proxy for actual land use files, giving our research team a general sense of the spatial-economic patterns in the city. Overall, the research showed that social media data, with all its biases, could be used to describe cities, where other data was not available.

Although technologically possible for the first time, using data from the public to describe cities this way is not new. Members of the Situationist movement of 1950's and 60's Paris believed the city should be understood through citizens' experiences as much as through traditional modes of representation, such as land use and

---

[35] In fact, these sites often sell this data to help businesses understand their popularity and overall position in the market.

[36] Sarah Williams, "*Here Now!* Social Media and the Psychological City," in *Inscribing a Square - Urban Data as Public Space*, eds. Katja Schechtner and Dietmar Offenhuber (Vienna, Austria: Springer, 2012).

census information.[37] Through the new field of Psychogeography, they postulated that only by understanding this experiential/emotional layer of the city could we begin to comprehend the city as a dynamic system.[38]

The Situationists were responding to the same concerns Jane Jacobs voiced in New York City - that governments had become too dependent on data as a mode of describing the city. These ideas relate to similar strategies of cognitive mapping popularized by Kevin Lynch, an urban theorist practicing in the 1960's and 70's. Cognitive mapping allows the public to mark important places in the city by drawing a mental map – which highlights the places that have meaning to them. Collectively, mental maps reveal public opinion on the importance of various city infrastructures. In many ways, geo-registered social media applications are tools for developing one's own cognitive map, as they allow users to express their personal experience in the city, thereby referencing the techniques employed by both the Situationists and Lynch.

---

[37] Denis Cosgrove, "Maps, Mapping Modernity: Art and Cartography in the Twentieth Century," *Imago Mundi* 57, no. 1 (2005): 35-54.

[38] Peter Wollen, "Mapping: Situationist And/or Conceptualist," in *Rewriting Conceptual Art*, ed. Michael Newman (London: Reaktion, 1999), 30.


FOURSQUARE CHECK INS - TOKYO

Image 4: Shows an image of Tokyo from the "We are Here Now" project. The transit-oriented development patterns in Tokyo become clear, with almost all businesses clustering around transit nodes. Image provided by Sarah Williams and Juan Francisco Saldarriaga.

VII. TEAMING UP TO WORK WITH DATA

Developing teams where each member specializes in a particular strategy to communicate with data creates the best results for using data to inform policy decisions. Teams benefit from working together to understand the complexity of the issue at hand, to visualize the information clearly, and to perform the technical tasks associated with data collection and analysis. Policy experts and political scientists identify policy questions, which can benefit from communication to larger publics. They also connect the group to a policy network, thereby providing access to data, other material resources, and an intellectual resource to critique the team's work. Data scientists excel at identifying both data and algorithms to answer civic questions, yet often do not know how to communicate the complexity of their results to broad audiences. This is why it is important to have designers as well as data scientists on teams. Graphic designers help edit the quantitative results into pieces people can visually consume. Technologists can identify new ways to collect and process data. Combined, these experts generate projects that transform data analysis into narratives that can communicate clearly and responsibility to broad audiences.

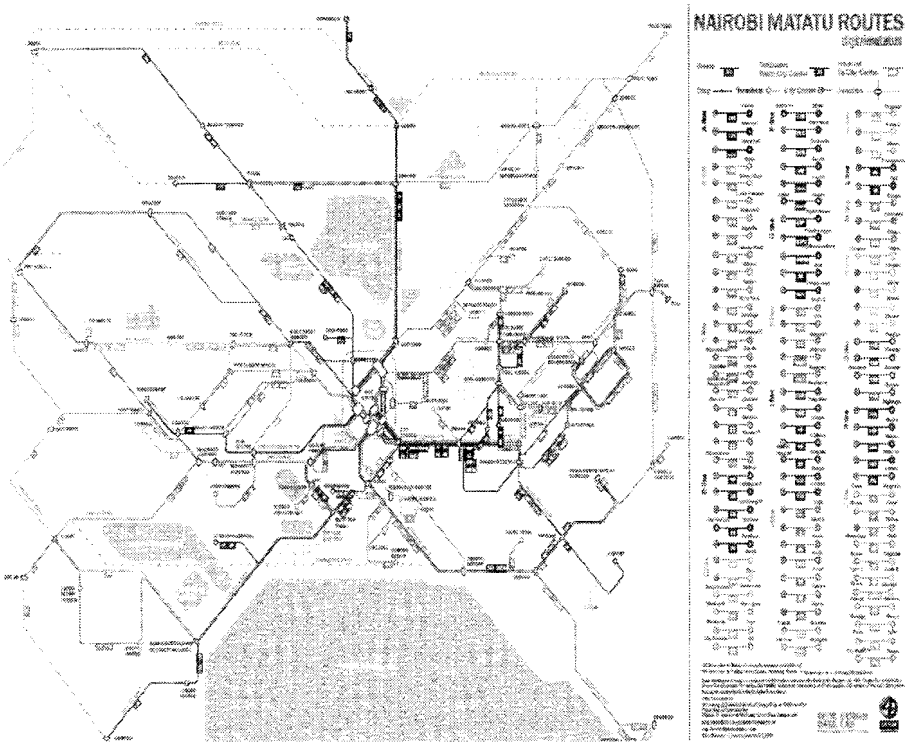

Image 5: Transit map for Nairobi. It was not until the Nairobi government could see the data in the form of a map that they were able to support the project. Image provided by Sarah Williams.

Digital Matatus, a project that uses mobile technology to collect data on Nairobi, Kenya's informal bus system, share that data in a standardized format.[39] Developing a transit map from that data exemplifies what it means to develop a team to make data actionable. The map was created through a partnership among the University of Nairobi's School of Computing and Informatics, Massachusetts Institute of Technology's Civic Data Design Lab, Columbia University's Center for Sustainable Urban Development (CSUD), and Groupshot, a technology-consulting firm with several projects in Nairobi. Each team member contributed his or her skill set to the

[39] Jacqueline Klopp, Sarah Williams, Peter Waiganjo, Daniel Orwa, and Adam White, "Leveraging Cellphones for Wayfinding and Journey Planning in Semi-formal Bus Systems: Lessons from Digital Matatus," in *Nairobi Planning Support Systems and Smart Cities*, eds. S. Geertman et al. (Cham, Switzerland: Springer International Publishing, 2015).

project.[40] The University of Nairobi team identified the matatu routes and developed mobile tools to collect the route data. They also built the technology in house so that the skills developed would remain in Nairobi. The Columbia University team, which was led by a political scientist with 20 years' experience working in Nairobi, helped connect the project to the political and institutional constructs of the transportation community in Nairobi. They developed workshops with the government, matatu owners and drivers, NGO's and the technology community to engage them in the creation of the data and map.[41] The interaction allowed the team to edit the map while also creating trust within the community that the data was valid.[42] MIT's Civic Data Design Lab developed the structural framework for the data collection, transformed that data into GIS files, and ultimately generated visualization strategies for the city's transit map. Groupshot helped connect the team to Nairobi's local technology community by promoting the existence of the data – which ultimately lead to the development of two transit routing applications. [43]

Overall the project was successful because we created a visualization that allowed everyone to see the transit data, engaged the local transit and technology community who accepted the data and built applications on top of it. The Nairobi government made it the official Matatu map and, at the map's official launch, started a conversation with the press about the state of transit in the city. The map and data have continued to be used by a UN-Habitat's project to plan a possible Bus Rapid Transit line along Thika Road in Naiorbi and by the matatu owners' organization to plan new routes in the city.

VIII. More Than Data

This essay has provided distinct strategies with examples for using data for civic change. When brought together, these tools can generate valuable experiments in civic action. Data visualizations are powerful as they bring numbers to life and deliver complex data analytics to the masses. Adapting mobile technology to create new data collection strategies allows individuals to describe their world

---

[40] Ibid.

[41] Ibid.

[42] Ibid.

[43] Interview with Sarah Williams, Director of the Civic Data Design Lab.

and push for new policies. Openly sharing data shifts the power dynamics inherent in the control of information, and allows anyone to use data to innovate, generating new products that increase civic potential. Narratives bring analysis to life by providing context for the patterns our data models expose. Digital traces are captured as we do almost anything, and social media data can provide a public voice to urban analysis. A comprehensive understanding of current policy questions, data analysis, and graphic techniques allows us to uncover new meanings hidden within data and make it actionable for civic change. Therefore it is essential to build teams that have skills in analysis, representation, and policy to make data relevant to larger audiences. Working with data on these multiple levels, reminding ourselves at every step the biases data can hold, helps to make data work for developing civic change.

In all the excitement about big data it is important to remember that data has been used inappropriately for civic interventions in the past, but it doesn't need to be used that way in the future. Data and its interpretation hold biases, and therefore it is important to read data representations through the eyes of those who create the information represented. Perhaps more importantly for those using data to make decisions, it is essential that their results be critiqued by the people it represents. The process of critique helps to include missing data and the voice of the public, thereby creating more balanced stories. Jane Jacobs argued that the voice of people must be included in data analysis that drives decisions. Technology is now allowing us to capture that voice. Whether it is citizen- sensing projects like the one described in Beijing or data gathered from social media sites, we now all have the ability to create cognitive maps. Incorporating quantitative data analysis with qualitative data that comes from the voices of people generates more balanced stories that help argue for policy change. Ultimately, the best way to ensure data is being used appropriately for civic change is to build teams of policy experts, graphic designers, and data scientists which allow data to be developed using multiple perspectives.