The Lightweight Smart City and Biases in Repurposed Big Data

Christian Voigt
Technology & Knowledge
Centre for Social Innovation
Vienna, Austria
voigt@zsi.at

Jonathan Bright
Oxford Internet Institute
University of Oxford
Oxford, UK
jonathan.bright@oii.ox.ac.uk

Abstract— **This paper addresses the implications of 'big data' on the smart city paradigm. In addition to grids of sensors to track traffic flows or monitor service delivery, urban governments around the world are starting to experiment with repurposing stores of data collected by third parties: using mobile phone data to track movement or social media to identify failing services. The use of this type of data has considerable potential to both augment the existing smart city vision and to spread it out to small and medium sized cities that are unable to afford investment in sensor grids, creating what we call a "lightweight" version of the smart city. However, it also implies a number of problems which previously smart cities were less prone to. After defining the lightweight smart city this paper reviews these challenges, mainly in the area of interpretation biases, before offering pointers to potential remedies and solutions.**

Keywords- ***Smart City; Big Data; Interpretation Biases.***

I. Introduction

Urban policymakers and planners are increasingly challenged by the scarcity of relevant and intelligible data, available in the policymaking contexts, particularly with the increased interest in accountability and transparency. The movement towards "smart cities" has often been presented as a way of fixing these problems. The smart city vision sees, as Kitchin puts it, "pervasive and ubiquitous computing and digitally instrumented devices built into the very fabric of urban environments" [1]. These devices promise a step change in the amount of data available to policymakers, and their corresponding ability to both create policy and respond to changing situations.

However, the smart city movement has been recently attracting more skepticism, for a variety of reasons. Some reports have highlighted the high up-front costs of installing large sensor grids, which in many cases seem to have been allied to relatively low returns [2]. These costs have also meant that, rather than spreading throughout the world, smart city technology is largely limited to a few urban megacities and one off projects, such as Songdo in South Korea and Masdar in Abu Dhabi. Furthermore, smart cities have been strongly criticized for promoting technological lock-in, by encouraging cities to sign large scale contracts with the IT services firms providing the infrastructure [1] [3], another factor which discourages investment on the part of governments. Finally, a variety of reports have critiqued the underlying focus of smart cities on business and enterprise, at the expense of other more progressive goals [4]. In this context, it is interesting to note the growing enthusiasm for "big data" within the smart cities movement. Big data is a concept which has attracted a variety of definitions [5], but for our purposes the key characteristic is that at least part of the definition involves a move to *creatively repurpose large stores of data which have been created as a by-product of another social activity*; for example, the use of Google query patterns to detect flu outbreaks [6], or Wikipedia search data to predict electoral outcomes [7] [8]. Big data are being drawn into a huge variety of fields and being used for a wide variety of different purposes. However, their use in the field of smart cities is particularly interesting: by offering the promise of relatively cheap, already collected data, they seem to provide a possibility for the smart city vision to break through some of the financial and technological barriers which currently impede it, and start being implemented around the world.

Our paper assumes a distinct social science perspective, as we focus on the societal implications that come with such a fundamental change in urban governance as the use of big data. Big data is effectively hailed as a game changer, turning classic hypothetico-deductive research into inductive analyses of big data [9]. Such grandiose statements try to establish a market for big data technologies from public, private-sector companies, such as IBM, Google, Facebook or Twitter. What is needed though, and increasingly delivered [10] [11], is a critical reflection on the inclusiveness of 'smart city' benefits, as well as a discussion of possible unintended effects, such as future dependencies in terms of data or technology lock-ins. It is useful to remember that cities have been trying to be 'data smart' before, e.g., using predictive computational models to address complex problems including city safety and public health in the 1960s [12]. But, as pointed out by Shelton et al. [11], "the fact that similar discourses are uncritically recycled by contemporary proponents of the smart city is troubling". The main aim of this article is to remedy this deficit, by discussing the potential implications of using *repurposed* big data in terms of information quality and potential interpretation biases. Its structure is guided by the following questions and thoughts:

- What difference could big data make in addressing some of the barriers to smart city adoption? We select prominent examples of smart city technologies and examine their potential from a social science perspective. On the basis of this discussion, we develop the concept of the "lightweight" smart city.

 ISBN: 978-1-61208-519-7

- What challenges might come with repurposed 'big data'? To avoid the trap of replacing old problems with new ones, we discuss some of the inherent challenges of governing cities by 'big data'.
- We conclude by opening up the discussion, suggesting a number of supportive activities, which make smart city services more accessible to an increasing number of cities and citizens.

The paper is organized as follows: section 2 introduces possible application scenarios for big data in smart cities. Then, section 3 gives an overview of known interpretation biases and their implications for lightweight smart cities. Finally, section 4 concludes with a discussion of measures to remedy distorting effects of interpretation biases and additional research needed.

II. HOW REPURPOSED BIG DATA AFFECTS THE SMART CITY VISION

As we describe above, using big data to drive smart cities involves enriching the vision (as described in [1]) of urban government using data provided by ubiquitous computing and sensor grids with the option of urban government making use of repurposed data coming from third parties, such as mobile phone companies and social media outlets [55]. In this section, we discuss the principal benefits of this move. The discussion is divided into three sections. First, we look at areas where repurposed big data can replace data generated by sensor grids. Second, we look at the use of big data to augment smart city technology (rather than replacing it), by optimizing the deployment of scarce resources and by providing new types of information. We conclude by arguing that big data offers the potential to provide a "lightweight" version of the smart city, which could potentially open up the smart city movement to a far greater range of cities, being less of a burden to already strained city budgets.

A. Sidestepping smart city sensor grids

The first way in which big data can support the smart city vision is in providing the potential for cheap data collection which does not require the installation of large scale sensor grids. Co-opting data from companies with stores of big data, such as mobile phone operators and social media providers is of course not cost free: license fees may need to be purchased, computing infrastructure may need to be set up to host the data, and skilled staff may be required to collect and process it. An example of this is provided by a recently completed collaborative study between Google and tw Netherlands Organisation for Applied Scientific Researchin Amsterdam [13]. They analyzed the extent to which anonymized urban mobility data from their Android mobile phone platform could be used to replace traffic sensor data on a 10 kilometer long stretch of highway. The results showed that the mobile phone data could duplicate the data provided by the sensors with high accuracy, "potentially saving €50,000 Euro per year [on that 10 km stretch of road alone] if the redundant sensors were removed". The potentially cheaper nature of data collection is allied to a second benefit, which is potential ease of implementation. To give an example of this, consider two different approaches to automatic failure detection in street lights, one found in Los Angeles and the other in the small town of Jun in Spain (which has just a few thousand inhabitants). Los Angeles has recently started rolling out smart LED street lighting along 4,500 miles of roads [14]. These lights communicate automatically with the bureau of street lighting, letting them know in particular if they are broken. This could be considered a classic implementation of part of the "smart city" vision: elements of the city themselves are able to communicate with government. The town of Jun, by contrast, has no such smart street lighting. However, what they have instead is a centralized effort to place the entire town on Twitter: everyone in city government and the vast majority of the residents have a Twitter account, and citizens are encouraged to interact with the government through this platform. The Huffington post gives an example of the way this works in practice [15], highlighting a case where a citizen noticed a streetlight had gone out, and sent a tweet to the mayor about the issue. The mayor responded that it will be fixed, with the Twitter handle of the engineer responsible also included, who himself tweeted the day after to notify that the streetlight had been fixed. Jun, in other words, have a kind of crowdsourced "smart" streetlight system [16] [17], with very rapid notification coming from citizens themselves.

B. Augmenting smart cities by optimizing resources and providing new data

Of course, there are many areas where repurposed big data will not be complete enough or accurate enough to fully replace smart city technology (or indeed parts of already existing government). However, in these cases, big data might still have a role to play in terms of optimising resources. For example, TomTom has recently started cooperating with Dutch police authorities, selling information about driver velocity from its Global Positioning System (GPS) tracking devices [18]. This information could not be used to directly convict people of speeding, both because it likely does not have the required degree of accuracy and also because of the concern TomTom itself would have to protect the privacy of its consumers. However, the authorities made use of the aggregate data to find the areas where speeding was most likely to occur, and then placed their mobile traffic cameras at these locations. In this case, big data does not replace the sensor grid, but rather augments it. Another example of this comes from the Mayor's Office of Data Analytics in New York [19]. One of the early successful projects this office worked on was a way to target restaurants which were illegally disposing of cooking oil into the city's sewers, something which was responsible for a considerable amount of blockages in the sewer system. The office compared data on restaurants which did not have an official oil disposal system with geographic information on sewer locations and blockages, in order to identify likely suspects of illegal dumping. These suspects were then visited by inspectors. Again, what this example shows is that this kind of big data technique does not replace existing information capture techniques used by cities. Rather, it augments them, allowing them to be directed more accurately and efficiently. Furthermore, there are also areas where big data driven smart cities go beyond its sensor driven counterpart. To characterize broadly, automatic sensors can be roughly classified into one of two types [20]–[22]. First there are sensors, which measure and report on characteristics of the physical environment, such as heat, light, the composition of the atmosphere, or the presence of physical objects. These types of sensors could, in a smart city context, provide real time indications of pollution, or measure water levels to check for flooding risks, automatically detect faults in lighting networks, etc. Second, there are sensors which not only report on the environment but try and capture data on the characteristics or behavior of people. However, there are still a great deal of policy relevant pieces of information smart city sensors cannot collect (i.e., which fall outside of these two types of sensor). This is where repurposed big data offers a chance to go further. For example, health problems are a key area of concern for policy makers. As is by now well known, Google has shown that it is capable of characterizing the size and duration of flu outbreaks from its search data [23], a result which has recently been extended to Wikipedia [24]; as well as other types of disease, such as dengue fever [25]. Another example would be the opinions and thoughts of citizens themselves on policy relevant topics, which a number of recent reports have flagged up as a potential source of information on policy specific topics, such as changes to a city's public transport system or opening a new shopping center [26] [27]. These examples demonstrate that big data offers a potential window into types of data which sensor driven smart cities could never hope to provide.

C. Towards a lightweight smart city?

In the terms that we have described them above, repurposed big data offers the potential for the implementation of a kind of “lightweight” version of the smart city. A lightweight smart city, based on repurposed big data, is like lightweight software in many respects. It is relatively cheap and easy to get going, requiring little special technological infrastructure to start up. Lightweight software is developed in order to increase the potential user base of the software: by making it easier to install and use, more people may take it up. Lightweight smart cities have similar potential consequences, potentially dramatically expanding the number of cities which can engage in “smart” programs. Thus far, almost all examples of smart city work come from large and economically powerful cities: in the UAE, in Singapore, in the US, in South Korea. These cities possess an obvious advantage for smart city work, which is that they have considerable budgets which can be put in to the creation of sensor grids. Small and medium scale cities are effectively shut out of the process. However, while offering much potential promise, the lightweight city also has an inherent potential challenge: the data being used within the city is no longer created or even owned by the city itself.

III. POTENTIAL BIASES IN BIG DATA FOR LIGHTWEIGHT SMART CITIES

In this section, we will move on to consider some of the challenges that a big data driven smart city faces, framed around the concept of bias. First off, we need to acknowledge that there is no agreed canon of terms and technologies, which constitute the 'smart cities' label. Hence, many criticisms to smart cities could possibly be discarded with reference to a different understanding of 'big' or 'smart' [9]. Smart city proponents claim that being empowered by new technologies (sensor enabled cars and streets, metered energy and water supply or people always connected and always tracked), governance is revolutionized, becoming more inclusive, performative and efficient [28]. Underlying these claims is a new paradigm of data-driven transparency or as New York's mayor Bloomberg is quoted "In God we trust. Everyone else, bring data." [29]. Although big data is probably as fuzzy a concept as smart city; the five Vs including volume, variety, velocity, veracity and value commonly describe big data [30]. Initially, there were only 3 Vs (volume, variety, velocity) and when the primarily technological challenges were solved, veracity and value was needed to justify the substantial investments made by smart cities [31]. However, we will argue that interpreting big data correctly and extracting value might be less straightforward than what we think. Biasing effects are a known phenomenon in information systems research, see [32] for a systematic overview. In general cognitive biases are not inherently detrimental to human judgment and decision making. Information filters, i.e., biasing the available information by not paying equal attention to all sources, are necessary mechanisms to deal with the constant influx of potentially useful data urban decision makers experience on a daily basis. Yet, Kahneman and Tversky showed that these filters are not always applied on a consistent and rational basis [33]. Depending on its presentation, the same data is perceived important or not (framing bias); similarly, data that is linked to recent events is more likely to influence people's decision making than data which is known to be important but has not had any recent appearance (recency bias). Hence, even though big data applications are meant to process vast amounts of heterogeneous data, this does not mean that biasing effects in designing and interpreting big data analyses would disappear. Jagadish [30] addresses a number of myths about big data including the misconception that big data automatically produces deep insights, without a need for theories. Multiple decisions are made, before big data analyses produce results. Following the big data life-cycle [30], these decisions concern *acquisition*, *cleaning*, *aggregation*, *modeling* and *interpretation* of data, decisions which in turn influence content, consistency and comprehensiveness of big data. However, the degree of comprehensiveness or consistency that can be realistically expected, depends on the problems big data analyses are applied to [34]. The following sections explore some examples of interpretation biases of mostly social media related big data.

A. Selection bias: How inclusive are the data sources?

Even though big data is generally said to be on the raise, access to big data might still hamper widespread analysis and research. Hence, the type of data cities might repurpose for their own uses can be limited. For example, control over the use of available data from most social media websites is restricted by service providers' business models, wherefore accessing large quantities tends to be either impossible or costly. An exception is Twitter, which allows users to access large parts of their historical data. Depending on the Twitter API (e.g., Twitter's freely available search or streaming APIs, or Twitter's commercial Firehose service) and the type of information requested, different amounts of Tweets can be acquired, ranging from tenth of thousand to several millions of tweets [35]. The availability of Twitter data has led to a number of studies investigating the use of Twitter as proxy for urban life or events impacting urban life. Nonetheless, prominent examples of Twitter's influence, such as the Arab Spring, the Obama elections or the Occupy Wall-street Movement, are often criticized due to a lack of systematic and more nuanced research [36]. Being aware of selection biases, we need to ask what a given set of big data is representing or suppressing, and whether our inferential claims are justified. For example, Arribas-Bel et al. [37] monitored geo-located Twitter activities (on average 1% of all tweets are geo-tagged) in order to understand activity levels in specific neighborhoods. The authors could show that activities in the virtual world of Tweets reflected expected behaviors as suggested by land use specifications (office space, residential area, tourism and leisure). In this instance, non-probabilistic sampling had been applied without drawing mistaken conclusions. However, there are questions about the inclusiveness of smart city data and their ability to represent elderly and economically isolated citizens [38]. Offenhuber reminds us that what citizens expect from smart cities is likely to be different depending on citizens' socio-economic status [39]. Citing the example of Boston where less affluent neighborhoods reported significantly less city maintenance issues through digital channels than areas that were better off. Offenhuber showed that there was not a lack of needs that prevented citizens from reporting more maintenance issues, but a mix of digital divide effect as well as a discomfort with calling on those who are accountable for city maintenance.

B. *Attentional bias: Are causations claimed where there are none?*

Whereas the neighborhoods analysis above discussed the issue of social media's representativeness of groups and activities in the physical city, Tufekci [40] highlights another issue which concerns the validity of conclusions drawn. Tufekci was observing social media used around Turkey's Gezi Park protests, exemplified by the use of the #jan25 hashtag. A frequency count over time showed a significant decline of the hashtag's use during June 2013. Concluding that the actual protest was declining in June, however, would have been far from correct, the topic became just so dominant that the hashtag was almost superfluous and was used less. This example is to illustrate that any data driven analysis might have blind spots, wherefore a theory is still needed, even though some big data proponents predict the end of theory as correlation supersedes causation [41]. The issue is magnified since with ever larger data sets, the likelihood of getting statistically significant results increases, leading to a proliferation of claims based on data patterns unrelated to the real world [42], also known as clustering illusion or the Texas sharpshooter fallacy [43].

C. *Framing bias: Does data interpretation reflect data collection?*

There is often an unstated assumption that 'hard' data is objective. Yet the matter of data is a matter of interpretation. Wilson [44] differentiates between the factual, representative side of data and its imaginative, urban-political side. In fact, as shown by the author, data can be used for diametrically opposed purposes. For example, citizens geo-mapped urban aspects, such as potholes or graffiti, which were simultaneously used to inform city officials about needed repairs as well as feeding into a 'desirable cities ranking' [44]. Clearly, whereas very active mapping would potentially lead to improvements of the build environment, it could also negatively impact the city's ranking and consequently the city's attractiveness for investors or a neighborhood's development prospects. Framing biases are also closely related to our assumptions about the nature of urban governance problems and the role scientific management and smart technologies can play. Criticism of prevalent 'Command-and-control' structures of IT-aided urban management in the 60s, highlighted already the inadequacy of cybernetic feedback loops, based on sensors, change actuators and controllers [45]. Goodspeed provides the example of urban renewal and freeway constructions and describes the situation as a wicked problem, one that has multiple, competing descriptions and where the solution requires value judgment and taking sides (i.e., land use decisions might create jobs and displace people at the same time) [45]. Clearly, there is no overriding single value that can be evoked in order to consent on the best decision. Hence, in such situations hard collective decisions need to precede the use of big data. The consideration of complex second and third order consequences cannot be delegated to big data if transparent decision making is a firm objective of smart cities.

D. *Information bias: Are some data more convenient than others?*

Information bias refers to the unwarranted over-interpretation of data; either through the way we classify, match and display data [46] or through including irrelevant data into their decision making and gain confidence where caution might be in order [33]. The classic example for the latter is Tversky and Kahneman's experiment of people ascribing jobs or study results to descriptions of people, containing little or no relevant information with regards to the question. The authors found that stereotypes, such as the clothing of librarians or a high degree of internal consistency of an in other ways completely irrelevant description of a person influenced people's confidence in their judgment considerably. Information bias can become a serious issue, when we think about predictive policing and the use of big data in law enforcement. New York City alone has 3,000 public surveillance cameras, 200 automatic license plate readers, 2,000 belt-mounted radiation sensors and diverse police databases [47]. This sensor driven city is then analyzed to identify high risk areas based on past crimes, but also circum-

 ISBN: 978-1-61208-519-7

stantial factors, such as text-mined tweets or Facebook postings related to specific areas [47]. As a consequence, these areas receive more police attention. Could big crime data replace human judgment in determining situations of reasonable suspicion, which would then lead to further investigations? Predicting citizens' behavior based on big data might represent a new privacy challenge, but as commented in [42], observational data are mostly generated and analyzed without citizens' knowledge and in public or open online spaces, where there is no right to be let alone. While privacy advocates call for a proper due process that ensures the right to be informed about how big data adjudicated a given course of action (police, land use permission, etc.) [48], others demand a more equitable distribution of riches made from user-generated content [42].

IV. Discussion

The smart city debate used to be about performance and competitiveness. Now we can arguably see a more inclusive debate emerge, addressing the reality of many small and medium-sized cities not being able to offer broadband in all city areas (let alone installing expensive sensor networks monitoring street lightening). Repurposed big data provides the potential for these cities to also innovate in the smart city debate. However, as shown in the previous section on biases, the lightweight smart city does not become automatically more inclusive by relying on smartphones and social media, mainly because these technologies are not equally distributed or used across all social groups. Still we think the benefits of the lightweight smart city outweigh the risks of biased interpretations. Being aware of the difference between data purposefully collected for urban management and repurposed, often social media-driven 'big data' is a first step to mitigate harmful effects of interpretation biases. Possibly biased data were also used prior to the raise of 'big data'. Simplified models of complex socio-economic systems can be as misguiding as uncritically following big data analyses. What is needed are dynamic interpretation and sense-making processes, that can supplement - rather than replace - urban management relying on traditional data sources, such as surveys, neighborhood meetings or purposeful observations. For repurposed 'big data' to be integrated successfully, urban management processes need to adhere to the proven principles of transparency and stakeholder participation.

Transparency. Early examples of master-planned cities have mostly reflected normative assumptions about 'good citizenship' [49], when in fact a substantial body of literature suggests that cities thrive on spontaneous encounters between people from all walks of life generating a collective creativity, which might get lost if people feel trapped in a virtual panopticon. Hence, cities need to be transparent about the data they collect and in what sense this data is repurposed. Even though reducing the opaqueness big data analytics (e.g., through required consultation and approval steps) can undo the efficiency gains also pursued by big data analytics [50]. Yet, more important than efficiency is the fairness of decisions - with or without the use of 'big data' -, wherefore citizens need a due process to question big data logic and, if necessary, police any unfair discrimination [51].

Participation. Greenfield smart city projects have shown that cities cannot be designed with technologies alone, smart cities understood as socio-technical challenges need socially rich innovation system, that enable learning, iterative experimentation and progressive social embedding of new technologies with existing stakeholders [49].

Limitations and future work. So far the paper has not covered a number of trends, such as the decreasing cost of sensors and wireless networking which avoids costly cabling [52]. Also, the emerging 'maker movement' might well enable cities to crowdsource data from citizens' sensors. Future research needs to explore the extend to which social media can support smaller cities if they do not have such a strong usage pattern as the Italian city of Jun or if privacy concerns motivate citizens to disable location tracking on their smart devices, so that it becomes increasingly difficult to geo-locate social media data. In the end, the question will be whether the lightweight smart city is better equipped than current variations of smart cities to address issues of economic growth as well as social inequalities.

Acknowledgment

The UrbanData2Decide project is funded under the Joint Programming Initiative Urban Europe (2014 -2016).

References

[1] R. Kitchin, “The real-time city? Big data and smart urbanism”, GeoJournal, vol. 79, no. 1, pp. 1–14, 2014.
[2] T. Saunders and P. Baeck, “Rethinking Smart Cities From The Ground Up | Nesta”, 2015.
[3] G. Graham, “Too-smart cities? Why these visions of utopia need an urgent reality check”, The Guardian, 13-Mar-2014.
[4] R. G. Hollands, “Will the real smart city please stand up?”, City, vol.12, no.3, pp. 303-320, 2008.
[5] V. Mayer-Schönberger and K. Cukier, Big Data: A Revolution that Will Transform how We Live, Work, and Think. Houghton Mifflin Harcourt, 2013.
[6] D. Lazer, R. Kennedy, G. King, and A. Vespignani, “The Parable of Google Flu: Traps in Big Data Analysis”, Science, vol. 343, no. 6167, pp. 1203–1205, 2014.
[7] T. Yasseri and J. Bright, “Predicting elections from online information flows: towards theoretically informed models”, [Available from: http://arxiv.org/abs/1505.01818]
[8] T. Yasseri and J. Bright, “Can electoral popularity be predicted using socially generated big data?”, it - Information Technology, vol. 56, no. 5, pp. 246–253, 2014.
[9] C. Rabari and M. Storper, “The digital skin of cities: urban theory and research in the age of the sensored and metered city, ubiquitous computing and big data”, Cambridge Journal of Regions, Economy and Society, vol. 8, pp.27-42, 2014.
[10] S. Abiteboul et al., “The elephant in the room: getting value from Big Data”, in Proceedings of the 18th International Workshop on Web and Databases, pp. 1–5, 2015.
[11] T. Shelton, M. Zook, and A. Wiig, “The 'actually existing smart city'”, Cambridge Journal of Regions, Economy and Society, vol.8, pp. 13-25, 2014.
[12] A. Townsend, “Cities of Data: Examining the New Urban Science”, Public Culture, vol. 27, no. 2 76, pp. 201–212, 2015.
[13] A. Eland, “Google Europe Blog: Tackling Urban Mobility with Technology”, 2015. [Online]. Available: http://googlepolicyeurope.blogspot.co.uk/2015/11/tackling-

urban-mobility-with-technology.html?m=1. [Accessed: 15-Mar-2016].
[14] A. Rogers, “8 Cities That Show You What the Future Will Look Like”, Wired Magazine, Available from: https://www.wired.com/2015/09/design-issue-future-of-cities/, [Accessed: 15-Mar-2016].
[15] W. Powers and D. Roy, “The Incredible Jun: A Town that Runs on Social Media”, Huffington Post, 20-Apr-2015.
[16] D. C. Brabham, Crowdsourcing, London: MIT Press, 2013.
[17] V. Lehdonvirta and J. Bright, “Crowdsourcing for Public Policy and Government”, Policy & Internet, vol. 7, no. 3, pp. 263–267, Sep. 2015.
[18] B. Waterfield, “Tom Tom sold driver”s GPS details to be used by police for speed traps”, The Telegraph, 28-Apr-2011.
[19] E. Copeland, “Big Data in the Big Apple: The lessons London can learn from New York's data-driven approach to smart cities”, Available from: http://capitalcityfoundation.london/wp-content/uploads/2015/06/Big-Data-in-the-Big-Apple.pdf, [Accessed: 15-Mar-2016].
[20] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman, “A taxonomy of wireless micro-sensor network models”, ACM SIGMOBILE Mobile Computing and Communications Review, vol. 6, no. 2, pp. 28–36, Apr. 2002.
[21] I. F. Akyildiz, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks”, IEEE Communications Magazine, vol. 40, no. 8, pp. 102–114, Aug. 2002.
[22] N. Xu, “A Survey of Sensor Network Applications”, IEEE Communications Magazine, pp.1-9, 2002
[23] J. Ginsberg et al., “Detecting influenza epidemics using search engine query data”, Nature, vol. 457, no. 7232, pp. 1012–1014, Feb. 2009.
[24] K. S. Hickmann et al., “Forecasting the 2013-2014 Influenza Season using Wikipedia”, vol.5, no.2, pp.1-15, 2015.
[25] N. Generous, G. Fairchild, A. Deshpande, S. Y. Del Valle, and R. Priedhorsky, “Global disease monitoring and forecasting with Wikipedia”, PLOS Computational Biology, vol.10, no.11, pp.1-16, 2014.
[26] “Social Media and Public Policy, September 2013 | The Alliance for Useful Evidence”. [Online]. Available: http://www.alliance4usefulevidence.org/publication/social-media/. [Accessed: 29-Feb-2016].
[27] “Use of social media for research and analysis - Publications - GOV.UK”. [Online]. Available: https://www.gov.uk/government/publications/use-of-social-media-for-research-and-analysis. [Accessed: 29-Feb-2016].
[28] C. Harrison et al., “Foundations for smarter cities”, IBM Journal of Research and Development, vol. 54, no. 4, pp. 1–16, 2010.
[29] L. Hoffmann, “Data mining meets city hall”, Communications of the ACM, vol. 55, no. 6, pp. 19–21, 2012.
[30] H. V. Jagadish, “Big data and science: Myths and reality”, Big Data Research, vol. 2, no. 2, pp. 49–52, 2015.
[31] E. Pinkerton and D. N. Edwards, “The elephant in the room: the hidden costs of leasing individual transferable fishing quotas”, Marine Policy, vol. 33, no. 4, pp. 707–713, 2009.
[32] P. J. Kirs, K. Pflughoeft, and G. Kroeck, “A process model cognitive biasing effects in information systems development and usage”, Information & Management, vol. 38, no. 3, pp. 153–165, 2001.
[33] A. Tversky and D. Kahneman, “Judgment under Uncertainty: Heuristics and Biases”, Science, vol. 185, no. 4157, pp. 1124–1131, 1974.
[34] R. Clarke, “Big data, big risks”, Information Systems Journal, vol. 26, no. 1, pp. 77–90, 2016.
[35] F. Morstatter, J. Pfeffer, H. Liu, and K. M. Carley, “Is the sample good enough? comparing data from Twitter”s streaming api with Twitter”s firehose”, arXiv preprint arXiv:1306.5204, 2013.
[36] Z. Tufekci and D. Freelon, “Introduction to the special issue on new media and social unrest”, American Behavioral Scientist, vol. 57, no.7, pp. 843-847, 2013.
[37] D. Arribas-Bel, K. Kourtit, P. Nijkamp, and J. Steenbruggen, “Cyber Cities: Social Media as a Tool for Understanding Cities”, Applied Spatial Analysis and Policy, vol. 8, no. 3, pp. 231–247, 2015.
[38] A. Glasmeier and S. Christopherson, “Thinking about smart cities”, Cambridge Journal of Regions, Economy and Society, vol. 8, no. 1, pp. 3–12, 2015.
[39] D. Offenhuber, “Infrastructure legibility—a comparative analysis of open311-based citizen feedback systems”, Cambridge Journal of Regions, Economy and Society, vol.8, no.1, pp. 93-112, 2015.
[40] Z. Tufekci, “Big questions for social media big data: Representativeness, validity and other methodological pitfalls”, arXiv preprint arXiv:1403.7400, 2014.
[41] C. Anderson, The end of theory: The data deluge makes the scientific method obsolete. Available: https://www.wired.com/2008/06/pb-theory/, 2008 [Accessed: Sept. 2016].
[42] H. Ekbia et al., “Big data, bigger dilemmas: A critical review”, Journal of the Association for Information Science and Technology, vol. 66, no. 8, pp. 1523–1545, 2015.
[43] H. D. Shane, Standard Deviations: Flawed Assumptions, Tortured Data, and Other Ways To Lie with Statistics, The Overlook Press: London, 2014.
[44] M. W. Wilson, “Data matter (s): legitimacy, coding, and qualifications-of-life”, Environment and Planning D: Society and Space, vol. 29, no. 5, pp. 857–872, 2011.
[45] R. Goodspeed, “Smart cities: moving beyond urban cybernetics to tackle wicked problems”, Cambridge Journal of Regions, Economy and Society, vol.8, pp. 79-92, 2014.
[46] P. Brey, “Values in Technology and Disclosive Computer Ethics”, in The Cambridge Handbook of Information and Computer Ethics, pp. 41–58, 2010.
[47] E. E. Joh, “Policing by Numbers: Big Data and the Fourth Amendment”, Washing. Law Review, vol.89, pp.35-68. 2014.
[48] K. Crawford and J. Schultz, “Big data and due process: Toward a framework to redress predictive privacy harms”, BCL Rev., vol. 55, p. 93, 2014.
[49] L. Carvalho, “Smart cities from scratch? A socio-technical perspective”, Cambridge Journal of Regions, Economy and Society, vol.8, no.1, pp. 43-60, 2014.
[50] T. Zarsky, “The Trouble with Algorithmic Decisions An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making”, Science, Technology & Human Values, vol. 41, no. 1, pp. 118–132, 2016.
[51] D. K. Citron and F. A. Pasquale, “The scored society: due process for automated predictions”, Washington Law Review, vol. 89, no.1, pp.1-33, 2014.
[52] G. P. Hancke and G. P. Hancke Jr, “The role of advanced sensing in smart cities”, Sensors, vol.13, no.1, pp. 393–425, 2012.

 ISBN: 978-1-61208-519-7