Big Data Solutions For Urban Environments

A Systematic Review

Francisco Ribeiro, Felipe Ferraz, Gustavo Alexandre
CESAR – Recife Center for Advanced Studies and Systems
Recife, Brazil
e-mail: {finr,fsf,ghsa}@cesar.org.br

Francisco Ribeiro, Felipe Ferraz,
Maria Silva, Gustavo Alexandre
Informatics Center
Federal University of Pernambuco
Recife, Brazil
e-mail: {finr, fsf3,mcts,ghsa}@cin.ufpe.br

Abstract— **Big data is a vast field of research and one that has not been completely explored. Studies involving big data, especially for solving problems in urban environments, still require more evidence. This work has the objective of identifying, evaluating, and interpreting published research that examine usages of the great amount of data generated by cities and its systems, in order to use Big Data technologies to improve city conditions. To achieve that, a systematic review of current literature was conducted. This review resulted in the finding of 1291 works of which 40 were identified as primary studies. The studies were then classified according to research focus and aspect of the city they focused on. The review investigates what it is known about the benefits and limitations in the use of big data in urban environments. The results show statistical data about big data, gaps in current research and models of successful implementation.**

Keywords— Big Data; Smart City, Systematic Review, Ubiquity

I. INTRODUCTION

Two global tendencies have been significantly influencing the information technology field in recent years: the growth in city population, marked by the migration of people from rural zones to urban ones; and the exponential rise of data generated by citizens stemming from the use of pervasive computing [1]. In this context, many challenges, opportunities for learning and for better management of cities arise through the use of new and extensive data sources, extracting data from them.

According to the United Nation Population Fund (UNFPA), 2008 was marked as the year in which more than half of the planets inhabitants, 3.3 billions, are now living in urban environments. With this unprecedented growth, many problems such as lacking transportation infrastructure, quality health assurance, citizens safety, unemployment rising, are driven cities in order to develop new and different means to try to avoid and mitigate those kind of issues, in order to provide citizens with a better living.

However, in tandem with the increase of the population also increases the amount of data that is generated by the same population. This can be leveraged to help improving city services and to better control and manage resources through the use of new technologies. Such as Big Data, data mining and/or data analysis.

The increase for data that is generated at a daily basis is explicit. According to Eron Kelly, “In the next five years we will generate more data as a human species than we have in the last 5.000 years” [2].

Develop and use new means to connect, integrate and to analyze such overwhelming amount of data is an important step to city continuity and competitiveness. It is vital to them to develop this kind of approach in order to help citizens providing them with a cohesive and intelligent view of the city.

The ubiquity of data sources and the resulting richness of information creates new research areas in computer and social sciences. Its application in diverse segments of industry is focus of many study efforts. However, the knowledge of the opportunities and challenges that arise through the use of big data technology applied to urbanization are still not sufficiently structured [1]. Albeit that many urban environments already use big quantities of data to improve infrastructure, planning and management, the term, as well as the underlying concept of big data are still lacking of more efforts. Among many different definitions, according to [4], big data refers not only to size of data but also to its speed and variety, and to obtain valuable information from this group of data, moreover, through the correct data analysis and leveraging on the size of the analyzed data it is possible to statistically increase the precision of the resulted information. In other words, a big data solution uses a group of data that is constantly increasing in order to provide its maintainers with more and more correct information that could be used to increase several aspects of a area.

The study contained herein is an effort to map out the current knowledge of the aforementioned issues. The conditions under which big data could be used strategically need to be understood, its limitations, anticipated and clarified.

This work is organized as such: In Section 2, basic concepts related to Big Data and the used methodology will be presented along with the objectives of the study. In Section 3, the methods, processes and the protocol that were used in the systematic review will be described. In Sections 4 and 5 the results related to the conducted research will be detailed presented. Finally, in Section 6 some conclusions and future works will be depicted.

II. BIG DATA

Big data is a term that is widely used in both academia and business. Despite the widespread adoption of the term, its meaning is still relatively unclear.

According to Cukier et al. [5], big data refers to the analysis of big quantities of data to find useful relationships and/or patterns. Other aspects can be seen in the definition of Michael Friedenberg – president and CEO of IDG Enterprise. According to him, the term refers to sets of data that are so big

 ISBN: 978-1-61208-445-9

-in terms of volume, speed and variety - that they become impossible to manage by conventional databases [6].

The amount of digital data that is currently being created on a daily basis by social network apps, embedded systems etc. is huge. Terabytes of information are produced from various types of sources

The application of big data technologies in urban environments can lead to truly intelligent cities that can be managed in real time with a high degree of precision [9].

III. APPLIED PROTOCOL

Based upon the guidelines for the development of systematic reviews in software engineering described by Kitchenham [10] and the analysis of the review model by Dybå et al. [11], a new methodology for revision was created. Our review methodology is composed of six steps: (1) development of the protocol, (2) identification of inclusion and exclusion criteria, (3) search for relevant studies, (4) critical assessment, (5) extraction of data, and (6) synthesis. The steps applied to the study contained herein are presented below:

The objective of this review is to identify primary studies that focus on the use of big data techniques that aim at solving urban problems. The following question helps identifying primary studies:

- *How is it possible to improve urban environment using big data and what are the challenges that accompanies the use of such technology for the creation of smart cities?*

From this central question, others secondary questions were developed of to help in the comprehension of the problem:

- *What aspects of urban environments can be optimized through the use of big data?*
- *Which solution models can the applied to cities?*
- *How can citizens and government officials benefit from the use of big data technologies?*
- *What are the main challenges in using big data in urban environments?*

A. Inclusion and Exclusion Criteria

For this review, we considered studies that aim at analyzing the use of big quantities of data to improve the efficiency and effectiveness of any indispensable city service. The studies could refer to specific sectors (e.g., traffic control or security) or have a broader scope taking into account many types of services. Since this field of research is recent, this review limited the examined studies to the ones published starting from year 2004.

Were also excluded:

- Studies not published in the English language;
- Studies that were unavailable online;
- Studies not based on research and that express only the official opinions of governments and field experts;
- Call for works, prefaces, conference annals, handouts, summaries, panels, interviews and news reports.

B. Search Strategies

The databases considered in the study is in the list below:

- ACM Digital Library;
- IEEE Xplore;
- ScienceDirect – Elsevier;
- SpringerLink.

Combinations of terms were created to guarantee that relevant information would not be excluded when querying different search engines and databases. As a result, four search strings were created:

1. "big data" AND (city OR cities);
2. "big data" AND citizen*;
3. "big data" AND urban*;
4. "big data" AND govern*;

In the process of extracting information from the databases, the search strings were used separately on each database. The searches were performed between December 2013 and January 2014. The results of each search were grouped together according to database and were, later, examined closer in order to identify duplicity. Table 1 shows the amount of studies found on each database.

TABLE I. AMOUNT OF STUDIES FOUND ON EACH DATABASE

Database	Number of studies
ACM Digital Library	396
IEEE Xplore	114
ScienceDirect – Elsevier	515
SpringerLink	290

C. Studies Selection Process

This Section describes the selection process from the beginning: from initial search using the Search Strategies described below to identification of primary studies.

At the first step, the studies that were obtained from the databases were gathered and added to Mendeley's citation management tool. This resulted in the finding of 1291 non-duplicated citations.

Secondly, the titles of all works selected in the previous step were analyzed to determine its relevance in this systematic review. At this stage, many works that did not mentioned using big data being to improve city conditions were eliminated.

Due to the use of terms related to city data, many works of geography, biology, medicine and sociology were found. In those cases, all works whose titles did not conform to the scope of the review were eliminated. In other cases, when the works titles were vague or unclear, they were put aside to be analyzed in the next step. At the end of this stage, 981 citations were excluded, thus remaining 310 items for further analysis.

In the third step, all abstracts of the works found in the

 ISBN: 978-1-61208-445-9

previous phase were assessed. Once more, many studies whose primary focus was not the analysis of large data directed to solving urban problems have been eliminated. We could conclude that the abstracts varied a lot in quality. Some items did not even had abstracts or had abstracts that did not clearly presented what the rest of the article was about. At this step, the articles that have not made it clear their conformity with the inclusion or exclusion criteria were included to be filtered out at a next step. Because of this phase, 251 studies were excluded, thus remaining 59 to be analyzed more closely.

Table 2 presents the amount of studies filtered in each step of selection process.

TABLE II. AMOUNT OF STUDIES FILTERED IN EACH STEP OF SELECTION PROCESS

Phase of Selection Process	Number of Studies
1. Databases Search	1291
2. Title Analysis	310
3. Abstract Analysis	59

D. Quality Assessment

In the quality assessment stage, works passed through a thorough critical analysis. In this stage, the complete studies were analyzed, instead of only the titles or abstracts. After this, the last studies that were considered uninteresting for the review were eliminated resulting in the final set of works. After the quality assessment, relevance grades were attributed to the remaining works. The relevance grades are going to be useful in the next stage.

Eight questions, based on [10] and [11], were used to help in the quality assessment. Those questions helped determining the relevance, rigor, and credibility of the work being analyzed. Of the eight, the first two are more useful to establish whether the work is relevant for the review, reason why they were the only ones used as exclusion criteria. The remaining questions are more useful for determining quality of the work, reason why they were used to grade the work according to quality. The questions were:

1. Does the study examine big data analysis as a way to aggregate value to citizens or rulers?
2. Is the study based on research - not merely on specialist's opinions?
3. Are the objectives of the study clearly stated?
4. Is the context of the study adequately described?
5. Were the methods for data gathering correctly used and described?
6. Was the research Project adequate to reach the research objectives?
7. Were the research results adequately validated?
8. Does the study contribute to research or citizens daily needs in any way?

Of the remaining 59 studies that were analyzed in the quality assessment stage, 40 passed to the stage of Data Extraction and Synthesis and were thus considered the primary studies. The quality assessment process will be presented in detail in the results Section along with the assessment of the 40 remaining studies.

IV. RESULTS

40 primary studies were identified [7], [12]–[50]. Each one deals with on a wide array of research topics and utilize a wide set of exploration models for each different scenario.

According to Ferraz et al. [51], it is possible to understand an intelligent city environment by interpreting six types of services: healthcare, transportation, education, security, government, and resources. Among the primary studies, we could find occurrences in four of the aforementioned groups. They were: healthcare, transportation, government and resources. The studies that did not fit into any of those groups were classified as being "general". The ones considered "general" are so for not relating to only one aspect of the city. They usually analyze the growth of data sources by focusing on the development of the city without taking into consideration the specific area they act upon.

A. Quantitative Analysis

The research process that was developed resulted in 40 primary studies. They were written by 124 authors linked to institutions based on 19 different countries, distributed on four continents, and were published between 2010 and 2013. In total, the authors identified 158 different keywords in their works.

In regards to the country of origin, most of the publications came from the United Stated (eight publications), followed by Germany, Greece, Italy, Australia and Spain (all with three works). China, Switzerland, Ireland and Korea was the third group (all with two publications). Each of the other remaining countries had only one publication. The large amounts of countries that have publications on the subject of big data show how widespread the topic is globally.

The most common keywords used in the remaining works with their respective frequency were: big data (10), smart city (8), data mining (4), social media (4), cloud computing (3), e-government (3), open data (3), open government (3), data fusion (2), innovation (2), personalized healthcare (2), twitter (2), ubiquitous computing (2). The first three keywords - big data, smart city and data mining - reflect precisely the theme of the research contained herein.

B. Quality Analysis

As was described in the previous Section, each of the primary studies was assessed according to eight quality criteria that relate to rigor and credibility as well as to relevance. If considered as a whole, these eight criteria provide a trustworthiness measure to the conclusions that a particular study can bring to the review. The classification for each of the criteria used a scale of positives and negatives.

Table 3 presents the results of the evaluation. Each row represents a primary work and the columns 'Q1' to 'Q8'

 ISBN: 978-1-61208-445-9

represent the 8 criteria defined by the questions used on quality assessment: Focus, Research, Objectives, Context, Data Gathering, Project, Validation, and Value, respectively. For each criteria, '1' represent the positive answer and '0' the negative one.

TABLE III. QUALITY ANALYSIS OF PRIMARY STUDIES

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Total
[7]	1	1	0	1	0	1	0	1	5
[12]	1	1	0	1	0	0	0	1	4
[13]	1	1	1	1	0	1	0	1	6
[14]	1	1	1	1	0	1	0	1	6
[15]	1	1	1	1	1	1	1	1	8
[16]	1	1	1	1	1	1	0	1	7
[17]	1	1	0	1	1	1	1	1	7
[18]	1	1	0	1	1	0	0	1	5
[19]	1	1	0	1	0	0	0	1	4
[20]	1	1	1	1	0	1	0	1	6
[21]	1	1	1	1	1	1	0	1	7
[22]	1	1	1	1	0	1	0	1	6
[23]	1	1	1	1	1	1	1	1	8
[24]	1	1	1	1	0	1	1	1	7
[25]	1	1	0	1	1	1	1	1	7
[26]	1	1	1	1	1	1	1	1	8
[27]	1	1	1	1	1	1	1	1	8
[28]	1	1	0	1	0	0	0	1	4
[29]	1	1	1	1	1	0	0	1	6
[30]	1	1	1	1	1	1	1	1	8
[31]	1	1	1	1	1	0	0	1	6
[32]	1	1	1	1	1	1	1	1	8
[33]	1	1	1	1	1	0	0	1	6
[34]	1	1	1	1	1	0	1	1	7
[35]	1	1	1	1	0	0	1	1	6
[36]	1	1	1	1	1	1	1	1	8
[37]	1	1	0	1	1	0	1	1	6
[38]	1	1	1	1	0	1	0	1	6
[39]	1	1	1	1	1	1	1	1	8
[40]	1	1	1	1	1	1	0	1	7
[41]	1	1	1	1	0	1	0	1	6
[42]	1	1	0	1	0	1	1	1	6
[43]	1	1	1	1	1	1	1	1	8
[44]	1	1	1	1	1	1	0	1	7
[45]	1	1	0	1	1	1	1	1	7
[46]	1	1	0	1	0	0	0	1	4
[47]	1	1	0	1	1	1	1	1	7
[48]	1	1	1	1	1	1	1	1	8
[49]	1	1	1	1	0	1	1	1	7
[50]	1	1	0	1	1	1	1	1	7
Total	40	40	27	40	25	29	21	40	

All studies that were analyzed in this step had positive answers for questions 1 and 2 because, as previously stated in the research methodology part, these questions represent inclusion and exclusion criteria. Consequentially, all studies with negatives answers to at least one of these criteria were already removed during selection stage.

All studies that were analyzed provided information on the context of the work and contributed in some way to research and the field development. A small part, 13 of 40 works, did not provide an objective description of the problem. A number of 15 works did not show properly its procedures for data collection and 11 works did not present its procedures for data analysis. 19 theoretical works did not provide validations for the proposed models or research results. Another fraction, 10 of 40 studies, obtained the maximum score in quality analysis. The highest number of negative answers was found for study 4.

V. DISCUSSION

After the analysis and data extraction steps performed on the primary works, it was possible to identify some aspects relating big data solutions for urban environments. In the first place, it was possible to conclude that big data applied to urban environments is a very recent field of research. All primary works were published after 2010. Second, a great quantity of theoretical solution and conceptual models are available. This shows that the research field is very active but also that concrete efforts to implement what was theorized are rare.

The review also showed that there is a lack of standardization in government solutions. This happens because unification of data policy is a hard task due to each government possessing specific priorities. In this context, solutions based on collaborative data, despite showing a great potential, need sustainable business models to ensure that all involved parties can profit.

A. *Urban sectors that can be optimized through the use of big data.*

After the analysis of primary studies, solutions could be categorized into four main urban sectors: government, healthcare, transportation and resources.

In the government sector, problems faced by both government officials and citizens were identified. Issues faced by government officials have to do with efficiency and effectiveness of organizational sectors [18][20]. Those issues are accentuated by urban population growth. Issues faced by citizens relate to lack of transparency of public sectors [16][38]. Citizens could benefit from accessing to government data because it would give them insight into how public money is spent.

In the healthcare sector the problem relates to high costs paid by government and general population. This happens as a result of fraud crimes [39] and inadequacies in medical tools. If doctors had better tools, diagnoses would be more precise and faster [22][29]. This would, in turn, reduce the amount of times the patients needed to return for consultations, benefiting the healthcare sector as a whole.

The main problem identified in transportation sector is the intense traffic that exists in cities [26][32][43]. This calls for alternatives in transportation to be developed and for constant monitoring of the traffic conditions.

In resources, we could verify that the difficulty lies in management. For efficient distribution, it is necessary that distribution networks be constantly being monitored in search for flaws and to aid planning of infrastructure [14][15][44].

B. *Solution Models.*

The solutions found in the revisions can be classified under the following categories: conceptual models, sensors and services, social networks and visualization techniques, and pattern recognition.

 ISBN: 978-1-61208-445-9

The conceptual models consist in revisions of the topic along with the development of theoretical models of recommendations based on research and learned lessons. This was the category that had the highest amount of solutions during revision where we can highlight the healthcare and government groups. Altogether, 19 of the analyzed studies are framed in this group: [7], [12]–[14], [16], [18]–[22], [27]–[29], [31], [33], [38], [40], [44] and [48].

The solutions based on sensors and services aggregate data from various sensors (be they heterogeneous or not) and provision them to be used in the development of services that can aid government officials or citizens. We identified 5 works at this group: [23], [24], [26], [36] and [49]. This type of solution, due to them not specifying what type of sensors can be utilized, can be applied in various areas of the city. Due that they were classified as being "general".

Solutions in social networks analyze data already available to infer tendencies, events, or reception of public policies. 4 studies were categorized at this class: [25], [34], [41] and [47]. Since this type of data can be analyzed with many different objectives, this type of solution can be found mostly in the "general" category.

The solutions in pattern recognition entailed mostly the development of many data analysis, machine learning and visualization techniques. The 12 works remaining were grouped at this category: [15], [17], [30], [32], [35], [37], [39], [42], [43], [45], [46] and [50]. They intend to extract relevant information from raw data and to present them in a way that can be used as a decision support tool. Many of those solutions were applied to problems of finding better routes to help in city traffic and were consequentially classified under the "transportation" category.

C. Value created by Big Data solutions

The application of big data technologies brings several benefits to urban sectors. In the "resources" category, the development of measurement structures used in tandem with data extraction technologies helped enormously in the management of water and energy. These technologies provide a number of advantages, including lower measurement costs, resource waste reduction for customers, theft detection, increased reliability of supply methods and the possibility of custom pricing strategies.

The benefits for the healthcare category vary from diagnoses that are more efficient to reduction in costs of medical systems. Most of the technologies that were revised are used to provide a more detailed analysis of patients' health by taking into account personal and medical history, and similarities to other patients to reduce the amount of medical consultations. Other benefits that medical systems provide stem from the analysis of data to reduce frauds, and other kinds of waste.

In transportation, the analysis of traffic patterns allows a more effective investment and provides the citizen with information on routes.

In government, most research highlights benefits of big data technologies for both citizens and managers. To citizens, the open government initiatives help providing more transparency in government spending, as well as allowing innovative solutions to be developed due to the availability of city data. To managers the benefits lie mostly in improvements of analytical works to aid in decision-making and predictions on impacts on society.

D. Towards Big Data challenges.

The main challenge in working with big data refers to how to deal with the overwhelming amount, speed and variety of that data. The size, speed and variety attributes turn traditional data analysis models obsolete or insufficient. Besides the problem of size, speed and variety, our study found another set of challenges:

- Standardization: The lack of standards in gathering and provisioning of data and lack of standards in solution models makes reuse difficult. As a result, solutions become very specific which can hinder the development on the field.
- Information Security: Most solutions use personal or sensitive data of people or businesses. The assurance of privacy, integrity and availability of that data is of the outmost importance when the objective is raising quality of life.
- Reversal and recuperation: Given the large amount of data and the vulnerabilities inherent to big data solutions, it is important to assure the recuperation of data in the case when data is lost due to some unexpected event. More intelligent solutions than Backup and Restore need to be created due to the large amount of data.
- Acceptance of Change: Some sectors are more resistant to the adoptions of new technologies – or of change in general. This happens often in government where people are accustomed to bureaucracy and to certain traditional ways of doing things.
- Quality: Data used by big data solutions often lack quality due to inexistent or insufficient validation. This is a challenge due to data validation being difficult when the size of it is very large. Errors can happened during gathering, analyzing or provisioning of data.

VI. CONCLUSION

The objective of this review was to identify studies and solutions that propose solving urban problems with application of big data. In the search phase, 1291 works were found of which 40 were classified as primary studies after selection and quality criteria were applied.

The studies were first classified in according to focus of their solutions. A big number of theoretical works and conceptual models were found. This showed that concretely implemented and validated solutions are currently scarce. Many solutions based on sensors and services were also found. Those solutions, however, lacked standardization. Other works presented data analysis models in social networks and pattern recognition techniques.

 ISBN: 978-1-61208-445-9

In regards to aspects of the city that are more frequently targeted in studies, we found the following: government, healthcare, transportation, and resources. Distinct solutions were identified for each and potential benefits were presented. The main challenges in concretely applying the solutions that were found are privacy, standardization, data quality, willingness to change, and security.

A limitation of the current study is the potential bias that exists when only one researcher is responsible for deciding selection criteria and analyzing quality of the works. Even though measures were taken to keep the analysis impartial, such as defining search questions and protocol beforehand, more strict validation, especially in regards to selection criteria, would benefit future research efforts. Another limitation is the lack of an in depth analysis of the solutions proposed by each study. The main focus of this work was to find patterns in the use of big data solutions applied to urban environments to provide a general view of the current state of art.

As a future work, more effort could go into the analysis of the solutions to enable the development of standardization models.

REFERENCES

[1] L. M. A. Bettencourt, "The Uses of Big Data in Cities," 2013.

[2] W. Redmond, "The Big Bang: How the Big Data Explosion Is Changing the World," 2012. [Online]. Available: https://www.microsoft.com/en-us/news/features/2013/feb13/02-11bigdata.aspx. [retrieved: January, 2015] .

[3] J. K. Laurila, J. Blom, O. Dousse, D. Gatica-perez, O. Bornet, and M. Miettinen, "The Mobile Data Challenge : Big Data for Mobile Computing Research." Proceedings of the Workshop on the Nokia Mobile Data Challenge, in Conjunction with the 10th International Conference on Pervasive Computing, pp. 1–8, 2012.

[4] S. Dirks and M. Keeling, "A vision of smarter cities," IBM, 2009.

[5] K. N. Cukier and V. Mayer-Schoenberger, "The Rise of Big Data. How It's Changing the Way We Think About the World.," Foreign Affairs, 2013. [Online]. Available: http://www.foreignaffairs.com/articles/139104/kenneth-neil-cukier-and-viktor-mayer-schoenberger/the-rise-of-big-data. [retrieved: February, 2015].

[6] M. Friedenberg, "Why Big Data Means a Big Year for Hadoop," 2012. [Online]. Available: http://www.cio.com/article/698750/Why_Big_Data_Means_a_Big_Year_for_Hadoop. [retrieved: February, 2015].

[7] R. Kitchin, "The real-time city? Big data and smart urbanism", GeoJournal, Volume 79, Issue 1 , pp 1-14, Nov. 2013.

[8] M. S. Fox, "City Data : Big , Open and Linked," pp. 1–20, 2013.

[9] M. Batty et al., "Smart cities of the future", Eur. Phys. J. Spec. Top., vol. 214, no. 1, Dec. 2012, pp. 481–518.

[10] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering", Technical Report EBSE-2007-01, 2007.

[11] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," Inf. Softw. Technol., vol. 50, no. 9–10, Aug. 2008, pp. 833–859.

[12] R. Nambiar, R. Bhardwaj, A. Sethi, and R. Vargheese, "A look at challenges and opportunities of Big Data analytics in healthcare," Big Data, 2013 IEEE Int. Conf., Oct. 2013, pp. 17–22.

[13] D. Arribas-Bel, "Accidental, open and everywhere: Emerging data sources for the understanding of cities," Appl. Geogr., Oct. 2013, pp. 1–9.

[14] D. Alahakoon and X. Yu, "Advanced analytics for harnessing the power of smart meter big data," Intell. Energy Syst. (IWIES), 2013 IEEE Int. Work., Nov. 2013, pp. 40–45.

[15] K. a. Nguyen, R. a. Stewart, and H. Zhang, "An intelligent pattern recognition model to automate the categorisation of residential water end-use events," Environ. Model. Softw., vol. 47, Sep. 2013, pp. 108–127.

[16] G. Lee and Y. H. Kwak, "An Open Government Maturity Model for social media-based public engagement," Gov. Inf. Q., vol. 29, no. 4, Oct. 2012, pp. 492–503.

[17] W. Radl, S. Jäger, F. Mödritscher, and A. Komendera, "And Data for All : On the Validity and Usefulness of Open Government Data," Proc. 13th Int. Conf. Knowl. Manag. Knowl. Technol., 2013, pp. 6–9.

[18] J. Bertot and H. Choi, "Big data and e-government: issues, policies, and recommendations," Proc. 14th Annu. Int. Conf. Digit. Gov. Res., 2013, pp. 1–10.

[19] H. Moon and H. S. Cho, "Big Data and Policy Design for Data Sovereignty: A Case Study on Copyright and CCL in South Korea," Soc. Comput. (SocialCom), 2013 Int. Conf., Sep. 2013, pp. 1026–1029.

[20] R. C. Joseph, P. State, and N. A. Johnson, "Big Data and Transformational Government," IT Prof., vol. 15, 2013, pp. 43–48.

[21] I. Vilajosana and J. Llosa, "Bootstrapping smart cities through a self-sustainable model based on big data flows," Commun. Mag. IEEE, vol. 51, no. June, 2013, pp. 128–134.

[22] N. V Chawla and D. a Davis, "Bringing big data to personalized healthcare: a patient-centered framework.," J. Gen. Intern. Med., vol. 28 Suppl 3, Sep. 2013, pp. S660–5.

[23] C.-H. Lee et al., "Building a generic platform for big sensor data application," Big Data, 2013 IEEE Int. Conf., Oct. 2013, pp. 94–102.

[24] F. J. Villanueva, M. J. Santofimia, D. Villa, J. Barba, and J. C. Lopez, "Civitas: The Smart City Middleware, from Sensors to Big Data," Innov. Mob. Internet Serv. Ubiquitous Comput. (IMIS), 2013 Seventh Int. Conf., no. 3, Jul. 2013, pp. 445–450.

[25] S. K. Bista, S. Nepal, and C. Paris, "Data Abstraction and Visualisation in Next Step: Experiences from a Government Services Delivery Trial," Big Data (BigData Congr. 2013 IEEE Int. Congr., Jun. 2013, pp. 263–270.

[26] L. Calderoni, D. Maio, and S. Rovis, "Deploying a network of smart cameras for traffic monitoring on a 'city kernel,'" Expert Syst. Appl., vol. 41, no. 2, Feb. 2014, pp. 502–507.

[27] W. Maass and U. Varshney, "Design and evaluation of Ubiquitous Information Systems and use in healthcare," Decis. Support Syst., vol. 54, no. 1, Dec. 2012, pp. 597–609.

[28] A. J. Jara, Y. Bocchi, and D. Genoud, "Determining Human Dynamics through the Internet of Things," Web Intell. Intell. Agent Technol. (IAT), 2013 IEEE/WIC/ACM Int. Jt. Conf., Nov. 2013, pp. 109–113.

[29] P. Bamidis, "Enabling e-services based on affective exergaming, social media and the semantic web: A multitude of projects serving the citizen-centric vision for ICT in support of pHealth," in Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on, 2013, pp. 1–6.

[30] C. Kaiser and A. Pozdnoukhov, "Enabling real-time city sensing with kernel stream oracles and MapReduce," Pervasive Mob. Comput., vol. 9, no. 5, Oct. 2013, pp. 708–721.

[31] K. Andreasson, J. Millard, and M. Snaprud, "Evolving e-government benchmarking to better cover technology development and emerging societal needs," Proc. 6th Int. Conf. Theory Pract. Electron. Gov. - ICEGOV '12, 2012, p. 430.

[32] E. Bouillet, B. Chen, C. Cooper, D. Dahlem, and O. Verscheure, "Fusing Traffic Sensor Data for Real-time Road Conditions," Proc. First Int. Work. Sens. Big Data Min. - SENSEMINE'13, 2013, pp. 1–6.

[33] S. Daniel, "geoSmartCity : geomatics contribution to the Smart City," Proc. 14th Annu. Int. Conf. Digit. Gov. Res., 2013, pp. 65–71.

[34] J. I. Criado, R. Sandoval-Almazan, and J. R. Gil-Garcia, "Government innovation through social media," Gov. Inf. Q., vol. 30, no. 4, Oct. 2013, pp. 319–326.

[35] A. Heise and F. Naumann, "Integrating open government data with stratosphere for more transparency," Web Semant. Sci. Serv. Agents World Wide Web, vol. 14, Jul. 2012, pp. 45–56.

[36] C. Dobre and F. Xhafa, "Intelligent services for Big Data science," Futur. Gener. Comput. Syst., Aug. 2013.

[37] X. Liu, F. Lu, H. Zhang, and P. Qiu, "Intersection delay estimation from floating car data via principal curves: a case study on Beijing's road network," Front. Earth Sci., vol. 7, no. 2, Jan. 2013, pp. 206–216.

 ISBN: 978-1-61208-445-9

[38] G. Puron-cid, J. R. Gil-garcia, L. F. Luna-reyes, and S. C. Martir, "IT-Enabled Policy Analysis : New Technologies , Sophisticated Analysis and Open Data for Better Government Decisions," Proc. 13th Annu. Int. Conf. Digit. Gov. Res., 2012, pp. 97–106.

[39] V. Chandola, S. R. Sukumar, and J. C. Schryver, "Knowledge discovery from massive healthcare claims data," Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '13, 2013, p. 1312.

[40] M. Bicking and M. Wimmer, "Need for computer-assisted qualitative data analysis in the strategic planning of e-government research," Proc. 11th Annu. Int. Digit. Gov. Res. Conf. Public Adm. Online Challenges Oppor., 2010, pp. 153–162.

[41] S. Nikolopoulos, S. Papadopoulos, and Y. Kompatsiaris, "Reality mining in urban space," in IISA 2013, 2013, pp. 1–4.

[42] A. Garzo, A. a. Benczur, C. I. Sidlo, D. Tahara, and E. F. Wyatt, "Real-time streaming mobility analytics," in 2013 IEEE International Conference on Big Data, 2013, pp. 697–702.

[43] A. Artikis, M. Weidlich, A. Gal, V. Kalogeraki, and D. Gunopulos, "Self-adaptive event recognition for intelligent transport management," Big Data, 2013 IEEE Int. Conf., Oct. 2013, pp. 319–325.

[44] D. S. Markovic, D. Zivkovic, I. Branovic, R. Popovic, and D. Cvetkovic, "Smart power grid and cloud computing," Renew. Sustain. Energy Rev., vol. 24, Aug. 2013, pp. 566–577.

[45] S. Hasan, C. M. Schneider, S. V. Ukkusuri, and M. C. González, "Spatiotemporal Patterns of Urban Human Mobility," J. Stat. Phys., vol. 151, no. 1–2, Dec. 2012, pp. 304–318.

[46] I. Carreras, S. Gabrielli, and D. Miorandi, "SUPERHUB: a user-centric perspective on sustainable urban mobility," Proc. 6th ACM Work. Next Gener. Mob. Comput. Dyn. Pers. Travel Plan., 2012, pp. 9–10.

[47] P. Metaxas and E. Mustafaraj, "The rise and the fall of a citizen reporter," Proc. 5th Annu. ACM Web Sci. Conf. - WebSci '13, 2013, pp. 248–257.

[48] J. H. Lee, M. G. Hancock, and M.-C. Hu, "Towards an effective framework for building smart cities: Lessons from Seoul and San Francisco," Technol. Forecast. Soc. Change, Oct. 2013.

[49] J. Yu and T. Zhu, "Towards Dynamic Resource Provisioning for Traffic Mining Service Cloud," Green Comput. Commun. (GreenCom), 2013 IEEE Internet Things, Aug. 2013, pp. 1296–1301.

[50] F. Giannotti, M. Nanni, D. Pedreschi, F. Pinelli, C. Renso, S. Rinzivillo, and R. Trasarti, "Unveiling the complexity of human mobility by querying and mining massive trajectory data," VLDB J., vol. 20, no. 5, Jul. 2011, pp. 695–719.

[51] F. S. Ferraz et al., "Towards a Smart City Security Model Exploring Smart Cities Elements Based on Nowadays Solutions", no. c, 2013, pp. 546–550.

 ISBN: 978-1-61208-445-9