SJPA
21(3)

Big Data in the Public Sector: A Systematic Literature Review

Cecilia Fredriksson, Farooq Mubarak, Marja Tuohimaa and Ming Zhan*

Cecilia Fredriksson
Faculty of Social Sciences, Business and Economy, Åbo Akademi University
cecilia.fredriksson@abo.fi

Farooq Mubarak
Department of Information Systems Science, University of Turku
farmub@utu.fi

Marja Tuohimaa
Turku School of Economics, University of Turku
marja.tuohimaa@gmail.com

Ming Zhan
Information Studies, Åbo Akademi University
mzhan@abo.fi

Keywords:
Big data
Public sector
Systematic literature review
Data management
Interdisciplinary research

Abstract

This systematic literature review describes how the concept of Big Data in the public sector is understood theoretically and practically in the academic field. The article's aim is to achieve a comprehensive understanding of Big Data in the public sector, an understanding which does not currently exist. In order to ensure an accurate investigation in accordance with the inclusion criteria, 156 items of literature were selected from 1,643 articles and publications. Five themes were selected in order to obtain more knowledge about the topic; evolution, definition, applications, challenges and the future of Big Data in the public sector. Our findings show that such themes have been emerging more widely since 2012, however, a definition of Big Data has still not been agreed. Big Data is most commonly described with reference to its characteristics. Practical applications related to Big Data in the public sector were consequently studied, as were the challenges and potential challenges that are emerging in the management of Big Data. There is great potential for public administration organisations to benefit from Big Data in the future, both from its practical operations as well as its strategic management in the public sector domain.

Introduction

The concept of Big Data is emerging as new challenges appear in analysing, archiving, sharing, transferring and processing large datasets across organisations. One reason for the growth of such huge numbers of datasets is digitalisation, the increased use of electronic devices and the growing popularity of social media (Hellerstein, 2008; Lohr, 2012). However, there is no uniform understanding of the concept within the academic community. Various definitions of Big Data exist and several attempts to describe Big Data have been made, but there is no agreement amongst scholars.

Various studies have explored the possibilities of Big Data, especially in private organisations (e.g. Chen, Tao, Wang, & Chen, 2015; Fan, Lau, & Zhao, 2015; Xiang, Schwartz, Gerdes, & Uysal, 2015), furthermore, scholars agree upon the value of using Big Data in the public sector. Manyika et al. (2011) state that the public sector can significantly boost productivity by actively using Big Data. According to Desouza and Jacob (2014), Big Data offers the potential to address

*** Cecilia Fredriksson** is a PhD student at Stockholm School of Economics and at Stockholm Center for Organizational Research (Score), in Stockholm, Sweden. She was a PhD student at the Faculty of Social Science, Business and Economics, Åbo Akademi University in Finland, when the article was written.

Farooq Mubarak is a PhD student at Turku School of Economics, University of Turku. Mubarak is writing his doctoral thesis about digital divide in developed and developing countries. His research interests include ICT, digital divide, and educational disparities between developed and developing countries.

Marja Tuohimaa is a Master of Economic Sciences from the University of Turku, Finland. Tuohimaa wrote her Master's thesis about a business intelligence model for strategic management in the public health care. She has been a manager within municipal organisations for ten years.

Ming Zhan is a PhD student at the Department of Information Studies, Faculty of Social Science, Business and Economics, Åbo Akademi University. Zhan is working on Big Data analytics concerning library-related posts on Instagram.

Scandinavian Journal of Public Administration
21(3): 39-61

ISSN: 2001-7405
e-ISSN: 2001-7413

many public sector problems, such as enhancing efficiency and productivity, transparency and well-being (Kim, Trimi, & Chung, 2014). Different digital methods are also seen as providing potential solutions to redress diminishing resources in public administration organisations (Kennedy, Moss, Birchall, & Moshonas, 2015). Podesta, Pritzker, Moniz, Holdren, and Zients (2014) state that Big Data is, in fact, already helping the United States Government in carrying out its obligations in respect of health, education and law enforcement.

Thus far, the rapid growth of data has attracted interest in how Big Data is used in the private sector (Manyika et al., 2011), leaving a gap in and a concomitant need for research into the public sector's use of Big Data (Desouza & Jacob, 2014). Expanding knowledge about the role Big Data plays in the public sector and how organisations can benefit from large datasets is thus necessary. Although increased attention has been paid to Big Data and its public sector role (see e.g. Mayer-Schönberger & Cukier, 2013; Murdoch & Detsky, 2013; Weiss & Zagroski, 2012), developing insight into the concept of Big Data in the public sector would also help to activate use of the unutilized data sets that currently exist in the public sector. This article consequently conducts a systematic literature review to study the concept of Big Data, specifically focusing on its use in the public sector. The review's main purpose is to achieve a comprehensive understanding of Big Data in the public sector, an understanding which is thus far lacking from the research field. To accomplish that aim, one main question, divided into two focus areas, is addressed:

How is Big Data understood in the context of the public sector
a) from a theoretical point of view, and
b) from a practical point of view?

To answer these questions and issues, a systematic literature review, giving an overview of Big Data in the context of the public sector, is conducted.

Methodology

A systematic literature review was chosen as the methodology to achieve the study aim. Compared to a traditional literature review, the advantage of this methodology is the rigorous filtering of relevant literature according to a predefined set of criteria. The literature chosen is therefore strictly relevant to the theme being examined.

Systematic literature reviews are considered to be original work because they are conducted following a rigorous, systematic process (Rother, 2007). They are performed in a manner that can be replicated, obtaining the same results with the same information. These reviews are different from traditional literature reviews because the review process is explicit and systematic (Khan, Kunz, Kleijnen, & Antes, 2003). There are five stages involved in the task, based on the steps in a systematic literature review proposed by (Khan et al., 2003). Firstly, the research questions were framed, with the relevant literature subse-

quently being identified. This involved defining keywords and databases with logical justifications. Next, the inclusion and exclusion criteria were established, according to which the references retrieved from online and manual searches were scanned and the relevant articles were marked. In step three, the quality of the literature retrieved was assessed and its relevance or not to the study was assessed. Step four involved an analysis of the content of the studies selected. In step five, the findings and recommendations of the study were presented.

In performing a systematic literature review, a careful selection of databases and keywords is important. Keeping this in mind, discussions and brainstorming were conducted among the authors, the resulting ideas were discussed with experts and the feedback resulted in the databases and keywords found in Table 1 and Table 2.

The databases that were selected were chosen because their literature (on business and the social sciences) matched our research objectives. Additionally, access to the full text of the article was requested and provided. Keywords were used in combination in the databases because one keyword alone would produce thousands of search results beyond the study's practical limits. The keywords were chosen because they cover the main scope of Big Data and the public sector. Since the targeted domain of the review is Big Data and the public sector, both were selected and then employed in combination in order to search the articles. In addition, according to Kitchin (2014), open data concerns all data types, especially those with the emphasis on public policy. Hence, studies of public sector open data were expected to indicate Big Data in that context. Open data within the public sector was consequently used for article collection as well. Based on the study by Xiang et al. (2015), data mining methods are conducted in order to extract the value of Big Data, i.e. data mining could be a solution for demonstrating the feasibility of Big Data and understanding its basis. Data mining and public sector were thus combined as keywords for article collection. In order to enlarge the scope of the search, municipality was selected as a key word in combination with Big Data. However, it was combined solely with Big Data rather than open data and data mining, so as to generate more related results. Additionally, Boiko (2005) considers content management a managerial operation regulating information from different resources – and because the public sector seems likely to involve several platforms for disseminating information, content management could be seen as a main public sector task. Big Data and content management were therefore, also combined as keywords.

Table 1: List of databases

Number	Database
1	EBSCO Host
2	Science Direct
3	Ebrary
4	Springer Link
5	Web of Science

Table 2: Keywords used for literature search

Number	Keywords
1	"Big Data" AND "public sector"
2	"Big Data" AND "content management"
3	"Data mining" AND "public sector"
4	"Open data" AND "public sector"
5	"Big Data" AND "municipality"

Table 3 shows the search results obtained from each particular database. The articles were scanned through defined inclusion and exclusion criteria to obtain the most relevant articles. This study did not focus on any particular country.

Table 4 lists the inclusion and exclusion criteria used for filtering out the relevant literature. There were 156 articles at this stage – attached as Appendix 1 to this paper. Although the inclusion and exclusion criteria were strictly observed, some additional references that were cited in the chosen articles were added during the process.

Table 3: Number of search results according to the different databases

Database	Total search results
EBSCO Host	211
Science Direct	730
Web of Science	68
Springer Link	610
Ebrary	24

Table 4: Inclusion and exclusion criteria

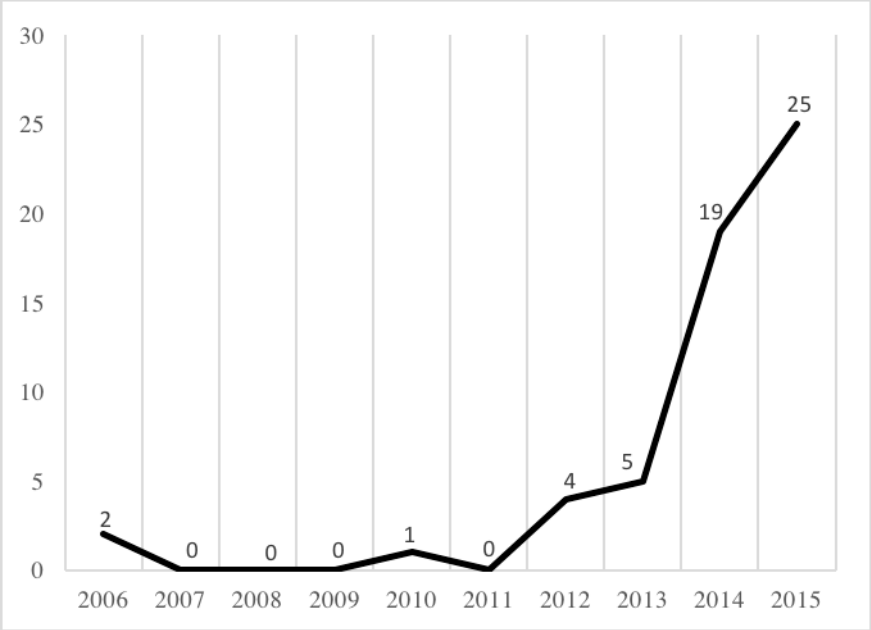
Inclusion criteria	Exclusion criteria
Articles in academic journals	Non-academic text
Empirical Studies	Interviews
Books	Book reviews
English language articles	Summaries of meetings or conferences
Fields of social science, information science, economic science, humanities	Field of agriculture science, technical science and health science
Theoretical studies	
Case studies	
Editorials	

These 156 articles were seen as the database for the content analysis. Every article was carefully read and the main focus of the article was thoroughly examined. To answer the research question and acquire a better result from the col-

lected article content analysis, the authors selected five themes. First of all, "definition" and "evolution" were chosen to answer the question of how the public sector understands Big Data in theory. Moreover, these aspects represent the past tense of Big Data in the public sector, which illustrates what has been achieved. "Application" and "challenge" aimed to answer questions concerning practical understanding, thus also representing the present and illustrating Big Data's current public sector situation. "Future" also concerns practical understanding in this systematic literature review, as it sheds light on Big Data's public sector development trends, which could help map the next stages of its practical development. Used chronologically those aspects usefully explain Big Data's present public sector situation, and were chosen as the basis for further analysis.

Ninety-nine articles were deleted from the list owing to an emphasis on business or technology issues rather than the public sector; articles with technology perspectives highlight detailed techniques or skills but do not understand the use of Big Data. According to the publication information of the 56 publications selected, 51 articles were found in journals (n=51) and only five came from books; all were published after 2000, see Figure 1.

Figure 1: Distribution of materials by year (n=56)



Since 2012, the number of publications in the target field has surged, indicating an especially prominent trend. The results of the content analysis of these 56 publications are presented with reference to the five aforementioned themes and reported in the next section.

The remainder of this article is divided as follows. The next section contains the content analysis results. They are presented according to the five themes selected. After that, the results are interpreted and the issue of how the public sector understands Big Data is discussed. The fourth section presents the conclusions and makes recommendations for future study.

The results of the content analysis

The evolution of Big Data in the public sector

With the development of information and communication technology (ICT), data generation, storage and processing has surged dramatically. Chen, Mao, and Liu (2014), note that advances in ICT make it easier to generate data and the fast development of cloud computing techniques is accelerating this process, plus such techniques can simplify data access and storage as well. According to Chen et al. (2014), ICT boosted the advent of Big Data. Under such circumstances, scientific paradigms have also evolved through empirical science, theoretical science and computational science and are now at the stage of data-intensive science (Chen & Zhang, 2014) – the latest approach to discovering knowledge or extracting value from the world. Big Data was not put forward as a term for creating more tools or technologies for data analysis until 2011 (Chen & Zhang, 2014). Since then, corporations, governments and scholars have devoted themselves to deriving meaning and value from Big Data.

Governments have a history of storing and codifying information for practical, legal and administrative usage (Henninger, 2013). Since the 1990s, information and ICT have been widely utilised by public agencies – initially originating from public administration departments. Governments have also tried to handle challenges in the recording and disposing of numerous data, but have experienced difficulties in managing the public and private sector while simultaneously performing crucial tasks (Liu & Yuan, 2015). Nonetheless, the public sector's attention to data occurred at least 10 years earlier than the ICT trend. In the late 1980s, some governments became concerned with the technical limitations of the data volumes, while in the 2000s, the integration of data was the main concern (Jordan, 2015) because a "data gold rush" appeared likely due to online activities, the declining cost of data management, widely used smart ICT applications and the continuous development of mobile communications (Kernaghan, 2014). Furthermore, web, text or opinion mining, online social networking, blogs, wikis, and forums have become ubiquitous and public sector actors now employ these tools to make decisions (Kő, Gábor, & Szabó, 2013). Meanwhile, the general public has become an active contributor to Big Data information by using social media. Consequently, the participatory paradigm of citizens in public decision-making processes has undergone great changes (Liu & Yuan, 2015) because citizens can now express opinions and requirements through social media; making it easier to participate in governmental activities. However, Big Data does not only exist in the field of decision-making, Wong et

al. (2015) note that Big Data has strong potential in the health care industry and encourages future research in this area. Meanwhile, Barack Obama's Open Government Movement (2009), has been copied in many countries, for instance, Australia, the United Kingdom and Italy (Veljković, Bogdanović-Dinić, & Stoimenov, 2014), boosting the concept of maximising the value of open data by reusing public sector information (Jung & Park, 2015).

The definition of Big Data

There is no consensus on Big Data's definition. According to Huang et al. (2015, p. 2), "there are many misunderstandings about Big Data" and one may be that any definition depends on the research angle (Kwon, Kwak, & Kim, 2015), nor has the context of Big Data been universally agreed on. Moreover, attention has focused on challenges and applications rather than Big Data per se.

Of the 56 publications selected, only 19 clearly defined Big Data, see Table 5:

Table 3: The core of Big Data in the different definitions

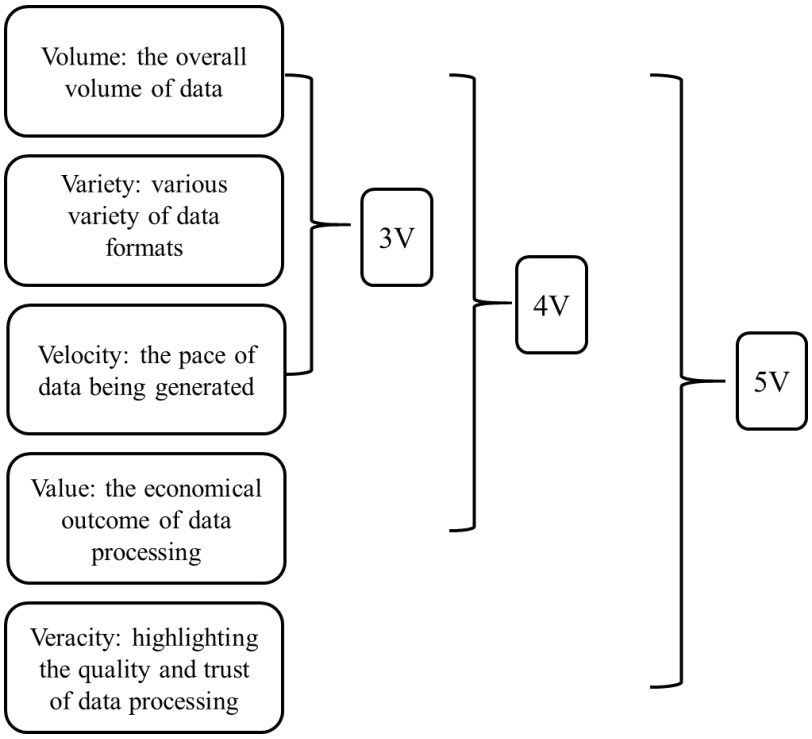
The core of Big Data	The Number of Articles
Data	6
Process	4
Features	9

This obviously demonstrates that defining Big Data is not a popular topic in current research. Three categories were created based on the main points emphasised in the 19 definitions. Six studies (Anna & Nikolay, 2015; Chen & Zhang, 2014; Hand, 2013; Kennedy et al., 2015; Srivathsan & Arjun, 2015; Stough & McBride, 2014) shared the understanding that Big Data is data (or databases) resulting in complicated issues for current technologies. In the six studies, Big Data was seen as a process that: 1) surpasses the ability of standard techniques to organise and manage it (Kennedy et al., 2015), 2) manages data by using the characteristics of the 5Vs (volume, variety, velocity, veracity and value) (Fosso Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015), 3) achieves new insights (Dobre & Xhafa, 2014) and 4) combines networks of technology, government, industry, markets, users and society (Shin, 2015).

According to Table 5, Big Data was mostly described by its prominent features, which are widely cited in the selected articles. Many scholars chose Three Vs (Volume, Velocity and Variety) to define Big Data (Jordan, 2014; Kitchin & Lauriault, 2014; Kshetri, 2014; McDermott & Turk, 2015; Pandey & Dhoundiyal, 2015). This idea stemmed from Laney (2001), who was first to suggest using three aspects to demonstrate data management: data volume, data velocity and data variety. Recently, practitioners have joined the discussion of Big Data resulting in Value being suggested as another characteristic for explaining Big Data (IDC, 2012; White, 2012), implying the potential value hidden in Big Data (Oracle, 2014). Consequently, Huang et al. (2015) chose Four Vs

(Volume, Velocity, Variety and Value) to define Big Data. Both the Three Vs and the Four Vs have been cited in other researchers' studies (Chen et al., 2014; Fosso Wamba et al., 2015). Based on former studies, White (2012) pinpointed Veracity as a new dimension to help express Big Data, regarding its data quality and as a result Fosso Wamba et al. (2015) used Five Vs (Volume, Velocity, Variety, Value and Veracity) to explain Big Data. The meaning of these Vs is explained in the Figure 2.

Figure 2: Five Vs



Applications of Big Data in the public sector

Although there is no consensus on Big Data per se, the benefits of Big Data applications are highlighted by many, e.g. the various governmental operations listed by Kim et al. (2014) prove the feasibility and the practical use of Big Data in the public sector. The following examples from selected literature illustrate how actors on different public sector levels use Big Data and what kind of purpose Big Data has. The overview is therefore presented based on two dimensions: the levels of the public sector (government actors, cities and citizens); and the functions (health care, social media and open data) Big Data has in the public sector.

Governments are central to both creating and managing knowledge. According to Jordan (2015), a government has two roles in knowledge work: the production of knowledge and the management of knowledge. Big Data has been seen as beneficial since it could actually allow governments to forecast policy demands, even before citizens express their opinions by voting (Jordan, 2014). Henninger (2013) also highlighted the importance of public sector information, especially in serving the public interest. Public sector information is used to inform government policy decisions, support economic development, increase citizen participation, generate trust in the government, and make knowledge accessible. Governments create information, by, e.g., collecting data from their citizens and institutions, and using that to support their own decision-making, planning and reporting (Henninger, 2013).

Empowered by the Internet of Things, *smart cities* aims at improving the quality of life of their citizens in different ways (Kamel Boulos & Al-Shorbaji, 2014). Numerous areas of local government are influenced by large amounts of data. For example, Alves, Martinez, and Viegas (2012) realised the significance of real-time data and suggested a new Intelligent Transport System, to provide information about, e.g. expected travel times. Different attempts to generate smart cities already exist, such as using information gained from website searches, visits and citizen feedback to create new knowledge about cities and offer new approaches to urban management (Arribas-Bel, 2014).

Attention is also focused on the individual – *the citizen* – as a central actor within different areas of the public sector (see e.g.Annoni, Ferrari, & Salini, 2006; Corallo et al., 2015). When discussing government data (Henninger, 2013), encouraging citizen engagement plays a key role in modern eGovernment, indicating data and public information are used to increase trust in governments and engage citizens in democracy.

The second dimension, the function of Big Data in the public sector, includes health care, social media and open data in this systematic literature review. According to Huang et al. (2015), Big Data technologies have many successful applications in different areas of *health care*. Researchers highlight several case studies that use Big Data in health care, e.g. health care records, recommendation systems, internet-based epidemic surveillance and using Big Data to determine air quality (Huang et al., 2015; Jordan, 2015). Furthermore, researchers can use secondary data sources – nationally required datasets from different statistics centres and surveys – to study disability and health (McDermott & Turk, 2015). Hence, Big Data has potential within the health care industry, encouraging future research in that subject area (Wong et al., 2015).

The use of *social media* is not only restricted to individual users, local governments benefit by using social media as decisions made at the municipal level have a direct impact on the daily lives of citizens. Hence, interaction between citizens and local authorities is deemed important, and the use of municipal webpages correlates with trust in local authorities (Lev-On & Steinfeld, 2015). Additionally, municipal Facebook pages contain significant potential for improving communication between citizens and authorities; for example, social media

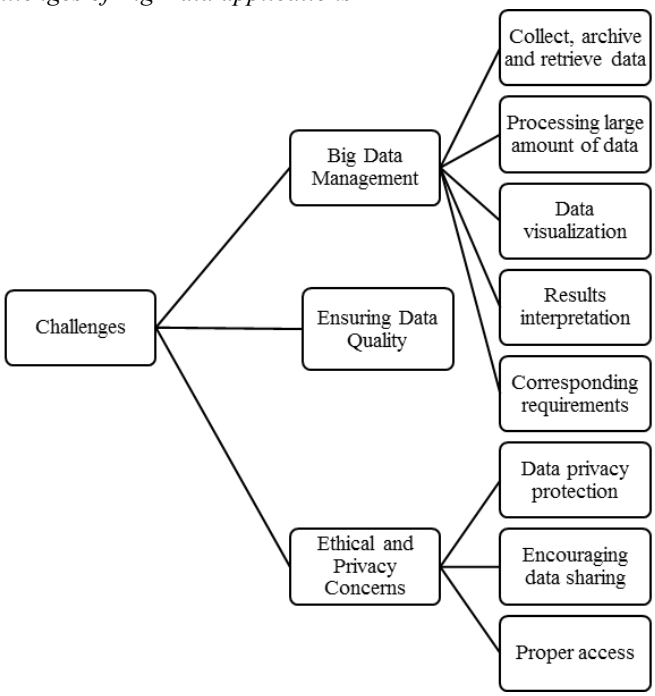
mining could provide a potential solution to the problem of declining resources and financial challenges within cities (Kennedy et al., 2015). As Kamel Boulos and Al-Shorbaji (2014) argue, the Internet of Things is providing new possibilities for improving smart cities.

Napoli and Karaganis (2010) emphasise that public policy should be made with publicly available data (or *open data*) to increase the level of transparency and its accessibility for decision-making; arguing that data transparency and access to it are fundamental to public policy making. Accordingly, there is a demand that governments are open and transparent, and, as stated by Henninger (2013), the benefits of using (open) government data are many – not only for the government itself, e.g. economic benefits, but for society as a whole, e.g. environmental impact benefits.

Challenges related to applying Big Data in the public sector

As shown, Big Data is said to generate opportunities for society. Nevertheless, its use also creates challenges, such as management challenges, data quality and ethical and privacy concerns, which all have the potential to occur due to the use of Big Data by the public sector. These challenges are summarised in Figure 3. The challenges are found through selected methodology and in the selected literature. However, although these challenges also exist on a general level outside the public sector, they are only discussed here with regard to the public sector.

Figure 3: Challenges of Big Data applications



Our systematic literature review reveals several management challenges associated with Big Data. These *Big Data management* challenges include data processing challenges, managing data across organisations and an absence of the technologies and expertise required to manage large amounts of data. Huge data volumes contain potential but their collection, archiving and retrieval is still a major challenge (Cao, 2012; Chen & Zhang, 2014; Dobre & Xhafa, 2014; Einav & Levin, 2014; Özköse, Arı, & Gencer, 2015). As noted by Raad, Al Bouna, and Chbeir (2015), it is difficult to handle the ever increasing volumes of Big Data. Likewise, Jiao et al. (2013) note that the characteristics of volume, variety and velocity create challenges for Big Data management. The vast amounts and the range of different data types as well as the high speed of data generation are all problematic when working with Big Data. Consequently, processing it all is challenging (Pandey & Dhoundiyal, 2015) and several methods such as data mining, data reduction and artificial intelligence have been proposed in order to cope with the challenges resulting from the continual generation of Big Data (Quick & Choo, 2014).

Large scale databases also lead to challenges in developing data visualisation (El Kadiri et al., 2015), common concerns include efficient analysis and visualisation (Chen et al., 2014). Regarding data analysis, using databases with other resources and interpreting the results with other materials are also issues for Big Data management (Arribas-Bel, 2014; El Kadiri et al., 2015). There are also challenges related to turning Big Data into valuable knowledge and understanding the outcomes of Big Data (Anna & Nikolay, 2015; Chen et al., 2014). According to Jordan (2015), the problem is not necessarily in the integration of the data, the problem lies is the availability, analysis, finding and understanding of relationships and turning it into usable knowledge. Big Data management requires expertise and skills if we are to extract useful insight from it (Kernaghan, 2014).

There are also challenges in archiving, distributing, and retrieving records across different platforms (Aljunid et al., 2012; Chen & Zhang, 2014). When dealing with data, problems in variation might occur because data has been merged from different sources. Depending on the dataset, a variation in its treatment or regional differences might cause problems for the generalisation of findings from different analyses and for the overall management of data (Hoagwood et al., 2015). To be able to manage increasing amounts of data, public sector actors need to develop technologies, management structures and programming capabilities (de Miranda Santo, Coelho, dos Santos, & Filho, 2006; Fuchs, Höpken, & Lexhagen, 2014). According to Kernaghan (2014), the development of Big Data technologies presents challenges for governments. Furthermore, Big Data technologies are needed to manage different data sources, e.g. the Internet and social media. As Liu and Yuan (2015) argue, the lack of new management approaches, governance structures and policy frameworks are obstacles to governments operating effectively. Napoli and Karaganis (2010) also note obstacles to gaining access to government data, such as regulations regarding commercially sensitive data, the confidentiality of regulators' internal delib-

erations and restrictive data licenses. Furthermore, unlike government agencies, third parties, such as public interest groups or industry organisations, are spared having to follow certain regulation regarding data management. Thus, further challenges and implications will likely arise for governments.

Lai and Hsiao (2014) stress the importance of the *quality of data*, and express concerns regarding actors, e.g. schools, having the capacity to collect and manage data in a way that produces high quality data. By quality they mean data that are appropriate for the purpose and provide sufficient evidence for decision-making. In health science, the quality of data needs to be considered since the merging of datasets is increasingly expanding and complex (McDermott & Turk, 2015). It is therefore important to clarify which datasets are comparable. Poor quality data can lead to unexpected and substantial costs for governments (Ohemeng & Ofosu-Adarkwa, 2015).

Privacy and security risks are also likely to correlate with the size, variety and complexity of data (Kshetri, 2014), thus it is understandable that questions regarding privacy and IT are a major part of the dialogue (Kernaghan, 2014) and there is concern among scholars regarding how data privacy is taken into consideration when managing Big Data (Terry, 2015; Truyens & Van Eecke, 2014). Encouraging individuals and other stakeholders to establish a trusted data sharing infrastructure is necessary, since privacy can be invaded during this process (Chen et al., 2014; El Kadiri et al., 2015). Similar challenges are also found in other studies (Batty et al., 2012; Fuchs et al., 2014). Critical opinions regarding the potential danger of public access to government datasets exist, such as the negative effects of a loss of public trust, which is often stated when discussing transparency and trust in governments (Henninger, 2013). Furthermore, there is a conflict between demands for privacy and those for more openness and transparency (Henninger, 2013).

Ultimately, the important questions about Big Data are likely to be ethical, not technical (Jordan, 2014). Furthermore, Jordan (2014) states that ethical knowledge is key, given the strong role that data analytics will play in public administration. Hence, IT's rapid development demonstrates the need for more research in the area of ethics and IT within the public sector (Kernaghan, 2014). There is a need for future research regarding Big Data and privacy and security risks, e.g. regarding differences in laws and regulations between countries (Kshetri, 2014).

Future of Big Data in the public sector

Research attention is also being paid to the field of Big Data and its future relevance (Chen & Zhang, 2014). Fan et al. (2015) point out several future study directions, such as selecting appropriate data sources and methods of analysis, dealing with heterogeneity among data sources, and refining frameworks as Big Data technology evolves. In addition, scholars suggest enriching Big Data future decision-making models (Daniell, Morton, & Ríos Insua, 2015; Fosso Wamba et al., 2015) while Fuchs and Horak (2008) examined real-time knowledge discovery as a future possibility. There is also concern that governments have not yet

done enough to allow open data to benefit the public. For example, government bodies and the industrial sector often make an issue of Big Data with regard to privacy, although that may result in poor decision-making (Wigan & Clarke, 2013). Furthermore, Stough and McBride (2014) recognise the growing importance of using Big Data across different disciplines and research institutions. The above studies draw attention to the shifting role of governments in providing open data to citizens and ethical concerns regarding data privacy and distribution (Sieber & Johnson, 2015; Washington, 2014).

From both user and service provider perspectives, Big Data contains several future opportunities. Shin (2015) for example, suggested that Big Data should be developed with a user-centric approach instead of traditional theory-based practices. Citizens could then take part in the Big Data revolution as there are numerous ways to share data (Wehn & Evers, 2015). Zhong, Huang, Müller Arisona, Schmitt, and Batty (2014) highlight the use of Big Data for the transport industry, through the use of data mining and analysis tools. Furthermore, the role of Big Data in health science is emphasized in research about achieving efficiency (Huang, Keser, Leland, & Shachat, 2003; Jordan, 2015). Jordan (2015) also expresses the importance of Big Data in health care by stating that applications created specifically for research involving access to health records would enable researchers to access all-encompassing health care data – if the patient approves. It is predicted that such data would generate insight and the ability to find trends and patterns, generating knowledge beneficial for customers and the market. Furthermore, Kennedy et al. (2015) argue that we need to find solutions to assist actors in the public sector who have limited access to the use of digital methods.

Correctly processed analyses may lead to many advantages for citizens. Data is constantly being generated in different forms, e.g. from social streams and mobile platforms, as content is scanned, recorded and archived (Teri, 2014). Public information is thus integral to economic development as it provides social services with information appropriate for delivering optimum services (Henninger, 2013). In the same way, Koerten and Veenswijk (2013) maintain that open access to public information is crucial to economic prosperity.

Discussion and interpretation

The motivation behind this review is to gain fundamental knowledge on how Big Data is understood in the public sector, which was employed as the lens for examining the selected articles. According to the result of the content analysis, Big Data is being applied in the public sector worldwide but is also generating challenges, which might direct its future development. To answer the research questions, the aspects of evolution and definition are discussed from a theoretical point of view, as these aspects represent the theoretical understanding of Big Data in the public sector. Furthermore, the aspects of application, challenges and the future are discussed from a practical point of view, as they mainly focus on the practical use and challenges related to the use of Big Data. The aspect of the

future shows the potential for practical developments to be made in the public sector.

Understanding Big Data in the public sector from a theoretical point of view

As a result of the methodology used, Big Data is conceptualised as a large volume of data with specific features, resulting in new insights and but also challenges for current technologies. Most scholars and practitioners would like to use the characteristics of Big Data to define it. To some extent, that might be helpful for explaining what Big Data is but these characteristics are fundamentally different to definitions. Definitions are rarely found in the empirical articles. One point is clear; there is no consensus on defining Big Data, which is reflected in the result of the three definitions in section 3.2. The fact that a clear definition of Big Data was found in only 19 of 56 studies echoes the study by De Mauro, Greco, and Grimaldi (2016), which also finds a lack of consensus on an agreed upon definition, something which might possibly be one reason for the lack of proposed definitions, rather than the methodology selected. Thus, we turn to literature reviews to shed light on a suitable definition as empirical studies mainly study the application of Big Data. In the context of the public sector, a definition of Big Data remains quite unclear but can be produced by referencing the definition of public sector information, which is "information, including all information products in any format, and services, generated, created, collected, processed, preserved, maintained, disseminated, or funded by or for public entities" (Henninger, 2013, p. p. 78). "Information products in any format" could be interpreted as "data in any type", which is a characteristic of Big Data, not to mention the worldwide expansion and explosion of information. The volume of such "information products" could be exponentially increased. Thus, Big Data is, in essence, reflected in public sector information. Additionally, "Open Data" – data which governments and other organisations make accessible to promote transparency and economic growth (Kotoulas, 2014) – are part of public sector Big Data. Although there is no common agreement on the concept or the definition of open data (Ohemeng & Ofosu-Adarkwa, 2015), there seems to be agreement about what constitutes open data. Open data is all information relating to public business in a democratic environment. In other words, the type, the amount and the growth speed of this information might have Big Data colours.

The development of ICT is a major reason for accelerating the advent of Big Data. Nonetheless, the public sector has been noting various types of data for a long time. Techniques, such as data mining, are generating possibilities for shedding light on data, thus enlarging the influence of Big Data in the private and public sectors. According to the selected articles, public decision-making forms the major part of this data-orientated evolution. Therefore, decision-making based on Big Data may become an indispensable process in almost every governmental department, revealing new opportunities and benefits. In the future, it is conceivable that open data and social media will be the aspects that drive Big Data as both are becoming popular in the public sector.

Understanding Big Data in the public sector from a practical point of view

Studies and statements regarding the application of Big Data in the public sector reveal a wide variety of examples of how state actors have used Big Data in practice to benefit government authorities and organisations, society and citizens. The areas in which Big Data and large amounts of data are applied vary from urban management to health care and citizen participation. For example, in urban management, Fuchs et al. (2014) introduced a knowledge infrastructure with the help of Big Data analytics, which is an intelligence-based information system for managing pre- and post-travel phrases. Data like website searching, booking and feedback are comprehensively used and the availability of such a system was evaluated in the city of Åre, Sweden. Similar examples can be found in other areas as well. There seems to be agreement on the significance of information flow and accessibility as well as the role it should have to improve transparency, enhance openness and increase citizen trust and engagement with government and its decisions. Ohemeng and Ofosu-Adarkwa (2015) express concerns about the imbalance between the government and civil society in creating an open data environment. There has been much focus on the role of the government, which can be seen in this review, nevertheless, scholars are repeatedly bringing forward citizen engagement as a key facet of the eGovernment and information society era. Different public-sector roles are important when discussing Big Data's future, hence, there is a request for greater focus on the role of civil society, rather than the role of the government. The variety of Big Data applications available to the public sector is also a sign of Big Data's importance to it.

The future of Big Data promises new developments, especially in the health care, industrial, and government sectors. Previous studies (e.g. Huang et al., 2015), on the potential for Big Data in the field of health care includes helping doctors to optimise time with patients and the quality of their service by enhancing the accuracy and speed of the processing of patient information through the use of correct and rapidly available information. In addition, Big Data can play a role in the establishing of electronic prescriptions via electronic communication between patient and health care provider. Another opportunity lies in a centralised health care system, which is possible through further research on Big Data and how to use it to develop sophisticated medical systems that connect pharmacies, health care providers and patients.

Apart from health care, more electronic government is possible, for example, allowing users to vote electronically and be active citizens. There are some initiatives in a number of countries but the scope is limited because not all citizens can participate in electronic governance. One clear advantage of an electronic voting system is that it allows transparency in its process. Moreover, development is required in electronic services to allow users to access such services via an easy to understand graphical user interface.

Big Data causes a diverse set of challenges and obstacles and many problematic issues exist for its management. Large amounts of varied data contribute to many challenges, such as the collecting and retrieving of Big Data and the

overall processing and visualisation of complex datasets. As stated by scholars, the governments and actors managing Big Data need to develop technologies, management structures and programming capabilities to be able to handle the exponential growth of data. Furthermore, concerns about Big Data's privacy and security risks are likely to increase as the amount of data increases. Thus, potential privacy and security risks need to be evaluated in all areas where Big Data is applied. Privacy will remain one of the largest Big Data challenges in the future as some people are reluctant to have their information kept in the form of electronic records, which is a major obstacle if relevant data is needed to centralise a digital service. Strict rules and regulations concerning data protection can also affect user perception of data privacy. Additionally, expert and skilled personnel need to be able to deal with existing challenges. Therefore, new management approaches, structures and policy frameworks are necessary to overcome the challenges identified above and transform Big Data into valuable knowledge and help understand the results of the analysis of Big Data.

Conclusions

This systematic literature review aims to achieve a comprehensive understanding of the use of Big Data in the context of the public sector, finding it to be an increasingly discussed concept in academia due to its potential for creating efficient processes for data management. The concept of Big Data in the public sector requires further research because the theoretical understanding of Big Data and how it is used in the public sector is currently insubstantial.

Even if the academic literature on the concept of Big Data is increasing, this systematic literature review shows that there is no agreed upon definition of Big Data among scholars. As an attempt to contribute to this gap, Big Data is here conceptualised as a large volume of data with specific features that can bring about new insights but which also challenges the abilities of current technologies to harness Big Data. A common description of Big Data is large volumes of data that are difficult to understand without the use of specific software and expertise. Furthermore, the findings show that a variety of characteristics, rather than definitions, are used to attempt to understand Big Data; hence, in the public sector context, the definition of Big Data is unclear. Nevertheless, public sector organisations and actors use Big Data. Furthermore, public administration organisations, cities and citizens produce and use Big Data to different extents. However, the outcomes of applying Big Data and open data in the public sector are considered similar because they are regarded as increasing efficiency and trust in government, while enhancing civic engagement and thus raising support for decisions made by state organisations. Big Data is also applied in different areas of the public sector, such as health care, urban management and citizen participation. The variety of applications that Big Data is used for in the public sector indicate its importance, motivating further research on it from a practical point of view, despite the many challenges associated with it.

A general picture of Big Data in the public sector has been illustrated in this review, and the selected publications signify the aspirations of governments to use it. Unfortunately, although many applications of Big Data have been launched, there are large knowledge gaps which require detailed attention and research in the context of the public administration of Big Data. Therefore, recommendations based on the findings of this literature review are made in order to outline potential future research areas.

Firstly, even though it is complicated, a universally shared definition of Big Data in the public sector context should be made in order to achieve a harmonised understanding. Furthermore, a clear definition could contribute to the harmonised use of Big Data in other sectors. Secondly, many applications have been produced with the idea of exploiting Big Data, but such applications rarely target governmental organisations, despite the public sector being composed of social organisations, like hospitals, universities, libraries – places where Big Data clearly exists. Therefore, examining Big Data applications for the public sector should be a future topic for researchers. Thirdly, the challenges associated with Big Data are widely recognised. Nevertheless, effective solutions for tackling specific problems exist. Problematically, whether the challenges are specific to certain situations or widely applicable whenever and wherever Big Data are employed is unclear. Determining whether such challenges are case specific or can be solved with off-the-peg answers would be a valuable area for future research.

To summarise, as a new area of study, Big Data provides possibilities and challenges. In particular, there are many complex issues to be considered when Big Data is studied in the public sector context and more studies are needed to provide solutions to aid its use by and for the people.

Acknowledgements

The authors would like to thank the Turku Urban Research Programme for funding the research project 'Big Cities meet Big Data' (2015-16), which have been a collaboration between Public Administration and Information Studies at Åbo Akademi University and the Department of Information Systems Science at University of Turku. This article is a result of this funding and collaboration.

References

Aljunid, Syed Mohamed & Srithamrongsawat, Samrit & Chen, Wen & Bae, Seung Jin & Pwu, Raoh-Fang & Ikeda, Shunya & Xu, Ling (2012) Health-Care Data Collecting, Sharing, and Using in Thailand, China Mainland, South Korea, Taiwan, Japan, and Malaysia, *Value in Health, 15* (1): 132-138.

Alves, David & Martinez, Luis M. & Viegas, José M. (2012) Retrieving Real-time Information to users in Public Transport Networks: An Application to the Lisbon Bus System, *Procedia - Social and Behavioral Sciences, 54* (2012): 470-482.

Anna, Kuraeva & Nikolay, Kazantsev (2015) Survey on Big Data Analytics in Public Sector of Russian Federation, *3rd International Conference on Information Technology and Quantitative Management, ITQM 2015, 55* (2015): 905-911.

Annoni, Paola & Ferrari, PierAlda & Salini, Silvia (2006) Data Mining Analysis on Italian Family Preferences and Expenditures in P. Perner (Ed.), *Advances in Data Mining. Applications in Medicine, Web Mining, Marketing, Image and Signal Mining*, Springer Berlin Heidelberg.

Arribas-Bel, Daniel (2014) Accidental, open and everywhere: Emerging data sources for the understanding of cities, *Applied Geography, 49* (May 2014): 45-53.

Batty, M. & Axhausen, K. W. & Giannotti, F. & Pozdnoukhov, A. & Bazzani, A. & Wachowicz, M., . . . Portugali, Y. (2012) Smart cities of the future, *The European Physical Journal Special Topics, 214* (1): 481-518.

Boiko, Bob (2005) *Content management bible*, John Wiley & Sons.

Cao, Longbing (2012) Social Security and Social Welfare Data Mining: An Overview, *IEEE Transactions on Systems, Man & Cybernetics: Part C - Applications & Reviews, 42* (6): 837-853.

Chen, C. L. Philip & Zhang, Chun-Yang (2014) Data-intensive applications, challenges, techniques and technologies: A survey on Big Data, *Information Sciences, 275* (10 August 2014): 314-347.

Chen, Jidong & Tao, Ye & Wang, Haoran & Chen, Tao (2015) Big data based fraud risk management at Alibaba, *The Journal of Finance and Data Science, 1* (1): 1-10.

Chen, Min & Mao, Shiwen & Liu, Yunhao (2014) Big Data: A Survey, *Mobile Networks and Applications, 19* (2): 171-209.

Corallo, Angelo & Fortunato, Laura & Matera, Marco & Alessi, Marco & Camillò, Alessio & Chetta, Valentina, . . . Storelli, Davide (2015) Sentiment Analysis for Government: An Optimized Approach in P. Perner (Ed.), *Machine Learning and Data Mining in Pattern Recognition*, Springer International Publishing.

Daniell, Katherine A. & Morton, Alec & Ríos Insua, David (2015) Policy analysis and policy analytics, *Annals of Operations Research, 236* (1): 1-13.

De Mauro, Andrea & Greco, Marco & Grimaldi, Michele (2016) A formal definition of Big Data based on its essential features, *Library Review, 65* (3): 122-135.

de Miranda Santo, Marcio & Coelho, Gilda Massari & dos Santos, Dalci Maria & Filho, Lélio Fellows (2006) Text mining as a valuable tool in foresight exercises: A study on nanotechnology, *Technological Forecasting and Social Change, 73* (8): 1013-1027.

Desouza, Kevin C. & Jacob, Benoy (2014) Big Data in the Public Sector: Lessons for Practitioners and Scholars, *Administration & Society*: 1-22.

Dobre, C. & Xhafa, F. (2014) Intelligent services for Big Data science, *Special Section: Innovative Methods and Algorithms for Advanced Data-Intensive*

ComputingSpecial Section: Semantics, Intelligent processing and services for big data, Special Section: Advances in Data-Intensive Modelling and SimulationSpecial Section: Hybrid Intelligence for Growing Internet and its Applications, 37: 267-281.

Einav, Liran & Levin, Jonathan (2014) Economics in the age of big data, *Science, 346* (6210): 715-721.

El Kadiri, Soumaya & Grabot, Bernard & Thoben, Klaus-Dieter & Hribernik, Karl & Emmanouilidis, Christos & von Cieminski, Gregor & Kiritsis, Dimitris (2015) Current trends on ICT technologies for enterprise information systems, *Computers in Industry*.

Fan, Shaokun & Lau, Raymond Y. K. & Zhao, J. Leon (2015) Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix, *Special Issue on Computation, Business, and Health Science, 2* (1): 28-32.

Fosso Wamba, Samuel & Akter, Shahriar & Edwards, Andrew & Chopin, Geoffrey & Gnanzou, Denis (2015) How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study, *International Journal of Production Economics, 165*: 234-246.

Fuchs, Christian & Horak, Eva (2008) Africa and the digital divide, *Telematics and Informatics, 25* (2): 99-116.

Fuchs, Matthias & Höpken, Wolfram & Lexhagen, Maria (2014) Big data analytics for knowledge generation in tourism destinations – A case from Sweden, *Journal of Destination Marketing & Management, 3* (4): 198-209.

Hand, DavidJ (2013) Data, Not Dogma: Big Data, Open Data, and the Opportunities Ahead in A. Tucker, F. Höppner, A. Siebes, & S. Swift (Eds.), *Advances in Intelligent Data Analysis XII*, Springer Berlin Heidelberg.

Hellerstein, Joe (2008) Parallel Programming in the Age of Big Data.

Henninger, Maureen (2013) The Value and Challenges of Public Sector Information, *Cosmopolitan Civil Societies: An Interdisciplinary Journal, 5* (3): 75-95.

Hoagwood, KimberlyEaton & Essock, Susan & Morrissey, Joseph & Libby, Anne & Donahue, Sheila & Druss, Benjamin, . . . Zerzan, Judy (2015) Use of Pooled State Administrative Data for Mental Health Services Research, *Administration and Policy in Mental Health and Mental Health Services Research*: 1-12.

Huang, Hai & Keser, Claudia & Leland, Jonathan W. & Shachat, Jason (2003) Trust, the Internet, and the digital divide, *IBM Systems Journal, 42* (3): 507.

Huang, Tao & Lan, Liang & Fang, Xuexian & An, Peng & Min, Junxia & Wang, Fudi (2015) Promises and Challenges of Big Data Computing in Health Sciences, *Special Issue on Computation, Business, and Health Science, 2* (1): 2-11.

IDC (2012) The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East.

Jiao, Yi & Wang, Yinghui & Zhang, Shaohua & Li, Yin & Yang, Baoming & Yuan, Lei (2013) A cloud approach to unified lifecycle data management in architecture, engineering, construction and facilities management: Integrating BIMs and SNS, *Advanced Engineering Informatics, 27* (2): 173-188.

Jordan, Les (2015) The problem with Big Data in Translational Medicine. A review of where we've been and the possibilities ahead, *Applied & Translational Genomics*.

Jordan, Sara R. (2014) Beneficence and the Expert Bureaucracy, *Public Integrity, 16* (4): 375-394.

Jung, Kyujin & Park, Han Woo (2015) A semantic (TRIZ) network analysis of South Korea's "Open Public Data" policy, *Government Information Quarterly*.

Kamel Boulos, Maged N. & Al-Shorbaji, Najeeb M. (2014) On the Internet of Things, smart cities and the WHO Healthy Cities, *International Journal of Health Geographics, 13* (1): 1-6.

Kennedy, Helen & Moss, Giles & Birchall, Christopher & Moshonas, Stylianos (2015) Balancing the potential and problems of digital methods through action research: methodological reflections, *Information, Communication & Society, 18* (2): 172-186.

Kernaghan, Kenneth (2014) Digital dilemmas: Values, ethics and information technology, *Canadian Public Administration, 57* (2): 295-317.

Khan, Khalid S. & Kunz, Regina & Kleijnen, Jos & Antes, Gerd (2003) Five steps to conducting a systematic review, *Journal of the Royal Society of Medicine, 96* (3): 118-121.

Kim, Gang-Hoon & Trimi, Silvana & Chung, Ji-Hyong (2014) Big-data applications in the government sector, *Communications of the ACM, 57* (3): 78-85.

Kitchin, Rob (2014) *The data revolution: Big data, open data, data infrastructures and their consequences*, Sage.

Kitchin, Rob & Lauriault, Tracey P. (2014) Small data in the era of big data, *GeoJournal, 80* (4): 1-13.

Koerten, Henk & Veenswijk, Marcel (2013) Public sector information reuse across Europe: Patterns in policy-making from an organizational perspective, *Journal of E-Governance, 36* (4): 198-211.

Kotoulas, Spyros (2014) Semantic and Reasoning Systems for Cities and Citizens in M. Koubarakis, G. Stamou, G. Stoilos, I. Horrocks, P. Kolaitis, G. Lausen, & G. Weikum (Eds.), *Reasoning Web. Reasoning on the Web in the Big Data Era*, Springer International Publishing.

Kshetri, Nir (2014) Big data′s impact on privacy, security and consumer welfare, *Telecommunications Policy, 38* (11): 1134-1145.

Kwon, Tae Hee & Kwak, Jeong Ho & Kim, Kiheung (2015) A study on the establishment of policies for the activation of a big data industry and prioritization of policies: Lessons from Korea, *Technological Forecasting and Social Change, 96*: 144-152.

Kő, Andrea & Gábor, András & Szabó, Zoltán (2013) Policy Making Improvement through Social Learning in A. Kő, C. Leitner, H. Leitold, & A. Prosser (Eds.), *Technology-Enabled Innovation for Democracy, Government and Governance*, Springer Berlin Heidelberg.

Lai, M. K. & Hsiao, S. (2014) Developing data collection and management systems for decision-making: What professional development is required?, *Studies in Educational Evaluation, 42*: 63-70.

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

Lev-On, Azi & Steinfeld, Nili (2015) Local engagement online: Municipal Facebook pages as hubs of interaction, *Government Information Quarterly, 32* (3): 299-307.

Liu, Shuhua Monica & Yuan, Qianli (2015) The Evolution of Information and Communication Technology in Public Administration, *Public Administration & Development, 35* (2): 140-151.

Lohr, Steve (2012). The Age of Big Data, *The New York Times*.

Manyika, James & Chui, Michael & Brown, Brad & Bughin, Jacques & Dobbs, Richard & Roxburgh, Charles & Hung Bayers, Angela (2011) *Big data: The next frontier for innovation, competition, and productivity*, Lexington, KY, McKinsey Global Institute.

Mayer-Schönberger, Viktor & Cukier, Kenneth (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Boston, New York, Houghton Mifflin Harcourt Publishing Company.

McDermott, Suzanne & Turk, Margaret A. (2015) What are the implications of the big data paradigm shift for disability and health?, *Disability and Health Journal, 8* (3): 303-304.

Murdoch, Travis B. & Detsky, Allan S. (2013) The inevitable application of big data to health care, *Jama, 309*: 1351–1352.

Napoli, Philip M. & Karaganis, Joe (2010) On making public policy with publicly available data: The case of U.S. communications policymaking, *Government Information Quarterly, 27* (4): 384-391.

Ohemeng, Frank L. K. & Ofosu-Adarkwa, Kwaku (2015) One way traffic: The open data initiative project and the need for an effective demand side initiative in Ghana, *Government Information Quarterly*.

Oracle (2014) Big Data for the Enterprise.

Pandey, Rajiv & Dhoundiyal, Manoj (2015) Quantitative Evaluation of Big Data Categorical Variables through R, *Proceedings of the International Conference on Information and Communication Technologies, ICICT 2014, 3-5 December 2014 at Bolgatty Palace & Island Resort, Kochi, India, 46*: 582-588.

Podesta, John & Pritzker, Penny & Moniz, Ernest J. & Holdren, John & Zients, Jeffrey (2014) *Big data: Seizing opportunities, preserving values*, Washington, The White House.

Quick, Darren & Choo, Kim-Kwang Raymond (2014) Impacts of increasing volume of digital forensic data: A survey and future research challenges, *Digital Investigation, 11* (4): 273-294.

Raad, Elie & Al Bouna, Bechara & Chbeir, Richard (2015) Preventing sensitive relationships disclosure for better social media preservation, *International Journal of Information Security*: 1-22.

Rother, Edna Terezinha (2007) Revisão sistemática X revisão narrativa, *Acta Paulista de Enfermagem, 20*: v-vi.

Shin, D.-H. (2015) Demystifying big data: Anatomy of big data developmental process, *Telecommunications Policy*.

Sieber, Renee E. & Johnson, Peter A. (2015) Civic open data at a crossroads: Dominant models and current challenges, *Government Information Quarterly, 32* (3): 308-315.

Srivathsan, M. & Arjun, K. Yogesh (2015) Health Monitoring System by Prognotive Computing Using Big Data Analytics, *3rd International Conference on Information Technology and Quantitative Management, ITQM 2015, 50*: 602-609.

Stough, Roger & McBride, Dennis (2014) Big Data and U.S. Public Policy, *Review of Policy Research, 31* (4): 339-342.

Teri, Tan (2014) Redefining and Repurposing Content as we Go Along, *Publishers Weekly*: 3-26.

Terry, Nicolas (2015) Navigating the Incoherence of Big Data Reform Proposals, *Journal of Law, Medicine & Ethics, 43* (1): 44-47.

Truyens, Maarten & Van Eecke, Patrick (2014) Legal aspects of text mining, *Computer Law & Security Review, 30* (2): 153-170.

Washington, Anne L. (2014) Government Information Policy in the Era of Big Data, *Review of Policy Research, 31* (4): 319-325.

Wehn, Uta & Evers, Jaap (2015) The social innovation potential of ICT-enabled citizen observatories to increase eParticipation in local flood risk management, *Technology in Society, 42*: 187-198.

Weiss & Zagroski (2012, 2012). Obama administration unveils “big data” initiative:announces $200 million in new r&d investments.

Veljković, Nataša & Bogdanović-Dinić, Sanja & Stoimenov, Leonid (2014) Benchmarking open government: An open data perspective, *Government Information Quarterly, 31* (2): 278-290.

White, Martin (2012) Digital Workplaces Vision and Reality, *Business Information Review, 29* (4): 205-214.

Wigan, Marcus R. & Clarke, Roger (2013) Big Data's Big Unintended Consequences, *Computer, 46* (6): 46-53.

Wong, Ho Ting & Yin, Qian & Guo, Ying Qi & Murray, Kristen & Zhou, Dong Hau & Slade, Diana (2015) Big data as a new approach in emergency medicine research, *Journal of Acute Disease*.

Xiang, Zheng & Schwartz, Zvi & Gerdes, John H. & Uysal, Muzaffer (2015) What can big data and text analytics tell us about hotel guest experience and satisfaction?, *International Journal of Hospitality Management, 44*: 120-130.

Zhong, Chen & Huang, Xianfeng & Müller Arisona, Stefan & Schmitt, Gerhard & Batty, Michael (2014) Inferring building functions from a probabilistic model using public transportation data, *Computers, Environment and Urban Systems, 48*: 124-137.

Özköse, Hakan & Arı, Emin Sertaç & Gencer, Cevriye (2015) Yesterday, Today and Tomorrow of Big Data, *Procedia - Social and Behavioral Sciences, 195*: 1042-1050.