Industrial Management & Data Systems

Safety or no safety in numbers? Governments, big data and public policy formulation

Joseph Amankwah-Amoah

Received 29 April 2015
Revised 29 July 2015
21 August 2015
Accepted 11 September 2015

Safety or no safety in numbers? Governments, big data and public policy formulation

Joseph Amankwah-Amoah
School of Economics, Finance and Management, University of Bristol, Bristol, UK

Abstract

Purpose – Although big data have emerged at the cornerstone of business and management research, past studies have failed to offer explanations and classifications of different levels of capacity and expertise possessed by different countries in utilising big data. The purpose of this paper is to examine the different capacities of governments in utilising big data.

Design/methodology/approach – The paper is based on a comprehensive synopsis of the literature on big data and the role of governments in utilising and harnessing big data.

Findings – The study provides an array of explanations to account for why some countries are adept at using big data to solve social problems, while others often faltered.

Research limitations/implications – The study offers a range of explanations and suggestions, which include skills upgrading, to help countries improve their capabilities in data collection and data analysis.

Originality/value – In this paper, data collection-data analysis matrix was developed to characterise the role of governments in data collection and analysis.

Keywords Capabilities, Public policy, Big data, Data

Paper type Research paper

Introduction

In less than a generation, the internet has transformed the way we live our lives and provided more opportunities for voluminous new data to be generated and disseminated across national borders (George *et al.*, 2014; Tapscott *et al.*, 2000). As more data are generated than ever before, big data have increasingly gained currency in the public policy arena (The Economist, 2010). Simultaneously, big data have taken one of the centre stages of contemporary discourse across the social science disciplines (Chen *et al.*, 2012). In recent years, however, governments around the globe have utilised big data to help combat contagious diseases (Milinovich *et al.*, 2015), identify potential terrorist and design and deliver effective health care (see Krumholz, 2014; Yiu, 2012; UK Department for Business, Innovation and Skills (DBIS), 2013).

In spite of accumulating evidence that suggests there are different roles and levels of capacity of governments in utilising big data (Brennan *et al.*, 2014; DBIS, 2013), much of the literature has failed to provide a clear framework to account for these differences. Whilst some governments are adept at using big data to solve social problems, others often falter and lack a clear path in utilising big data (see Kim *et al.*, 2014; The Economist, 2010). Scholars have overlooked this important issue and failed to offer robust explanations to account for these differences.

The main purpose in this study is to examine the different capacities and capabilities of governments in harnessing and utilising big data. This study seeks to address this gap in the literature by developing a 2×2 matrix to shed light on why some governments are better than others at utilising data and generating knowledge.

Industrial Management & Data Systems
Vol. 115 No. 9, 2015
pp. 1596-1603

0263-5577
DOI 10.1108/IMDS-04-2015-0158

The remainder of this paper proceeds as follows. The paper begins by presenting a brief overview of the literature on big data and then sets out the key feature of the framework. Based on this, we tease out the key features of our 2×2 framework. The paper concludes by highlighting a number of implications for theory and practice.

Governments and big data: towards a typology

Over the past few years, many organisations and governments have engaged in amassing, analysing and profiting from big data (Woerner and Wixom, 2015; Economist Intelligence Unit (EIU), 2011). Big data can be defined as the "ultralarge bodies of data that have not been prospectively limited in size or scope by the intent to address specific research questions" (Ghani *et al.*, 2014, p. 976). Big data entail the process of collecting, analysing and storing of data created.

Nevertheless, others view big data as "a troubling manifestation of Big Brother, enabling invasions of privacy, decreased civil freedoms, and increased state and corporate control" (Boyd and Crawford, 2012, p. 664). It is worth noting that democratic systems of governments are generally characterised by limited government and high private property protection, which limit ability to collect and analyse data on a large scale (see Cavusgil *et al.*, 2012; Amankwah-Amoah, 2015a; Millward, 2005).

An accumulating body of research suggests that governments are often confronted with the need to collect and analyse data across array of sectors such as health, education and economy (Brennan *et al.*, 2014; Kim *et al.*, 2014). These data are often stored and utilised across government agencies to guide public policy decisions. Accordingly, data collection and analysis play pivotal roles in knowledge generation geared towards public policy formulation (Mayer-Schönberger and Cukier, 2013).

However, past studies have largely failed to conceptualise the different capacities and capacities across countries. Building on Forte's (1994) and Ward *et al.*'s (1986) works on information and data, and the EIU's (2011) notion of "stages of evolution", we develop a data collection-data analysis matrix to classify governments relative to big data. We begin by contending that governments/countries, as a matter of routine, engage in two important activities, i.e. data collection and data analysis. It is also our contention that governments also possess different levels of expertise and capacities in these two areas which then shape their ability to extract maximum value from big data. Some governments engage in large-scale data collection whilst others do not. Therefore, countries/governments have a range of expertise from "data poor" to "data rich". Put another way, data richness and information richness are both viewed as a continuum (see Figure 1).

By data rich, we are referring to the mere possession of large data sets collected through formal and informal channels at various branches of government (see Ward *et al.*, 1986). The emergence of the internet and social media has meant big data are now readily available to be tapped into. By information rich, we are referring to the abilities and capabilities in analysing and converting big data into knowledge and understanding which then provide the foundation for public policy formulations and actions (Forte, 1994). This process also entails making use of business analytics, i.e. "the techniques, technologies, systems, practices, methodologies, and applications" employed to analyse big data (Chen *et al.*, 2012, p. 1166). Therefore, data analysis is needed to enrich understanding and inform our actions.

Recent years have witnessed a stream of research indicating that "for data to be informative it has to be summarised, given a useful context and acted upon, and it is here that decision support systems can make a valuable contribution to the

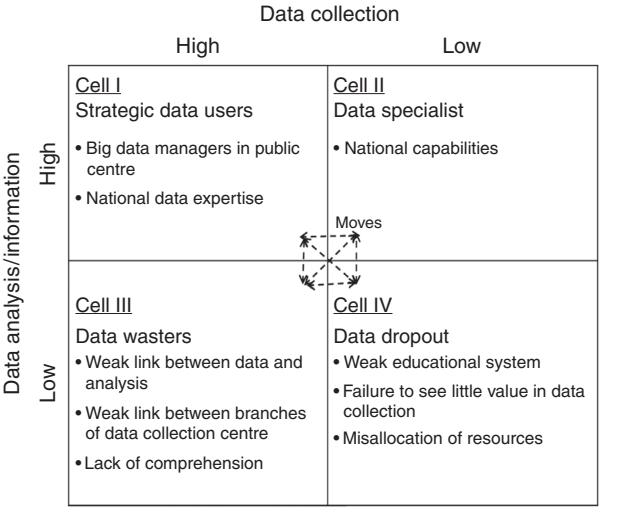

Figure 1.
Data collection-data analysis matrix

decision-making process" (Forte, 1994, p. 149). Big data capabilities include the ability to identify and interpret data to help inform public policy. It is worth noting that the mere possession of big data is unlikely to enhance a country's national competitiveness.

In this fast-changing world, national capability in big data is partly rooted in the ability to tap into reservoirs of data across government departments to identify new uses of old and new data. However, countries' ability to harness big data to solve social problems also depends on factors such as the availability of a skilled workforce and well-developed educational system (DBIS, 2013). Indeed, past studies have long established that "poor countries differ from rich in having fewer institutions to certify quality, enforce standards and performance, and gather and disseminate information needed" (WDR, 1998, p. 1; Amankwah-Amoah, 2014a, b; Amankwah-Amoah and Debrah, 2014).

To ground the framework on a stronger theoretical underpinning, data collection and data analysis are viewed as two key pillars. Crossing the two dimensions with a different degree of expertise produces the 2×2 matrix which classifies governments/countries, as demonstrated in Figure 1. Each of these cells emphasises different levels of capacity of national government which then influences their ability to utilise big data. Countries are able to move across cells.

Data sources

Before articulating the details the four cells, this section sheds light on the approaches to assemble data to inform the analysis. The present research draws from archival sources. Indeed, archival data has been found to be effective in theory development and conceptualisation (see Welch, 2000). The key archival sources include the World Development Reports and Economist Intelligence Unit Reports. In addition, reports on big data by the White House and UK Department for Business, Innovation and Skills were also examined. These sources provided further insights on how governments accumulate, analyse and utilise data.

Cell I: strategic data users (data rich, information rich)

This is where there is a very clear link between data collection, analysis and public policy formulation. "Strategic data users" are also more likely to process data across branches of government with pace and extract maximum value (see EIU, 2012). This can be attributed to advances in technologies that allow large data to be collected and analysed from sources such as the web, mobiles and social media to inform policy. This cell is also characterised by data sharing across government agencies, service providers and other branches of governments. Although globalisation is helping to close the knowledge gaps that exist between developed and underdeveloped countries, many developing countries still lag behind their developed country rivals in their ability to turn data into information and knowledge (see WDR, 1998; DBIS, 2013).

One possible explanation for this is that advanced economies generally have more effective institutions such as governments, regulations and laws for establishing and enforcing high standards which enable them to capture data, analyse it and turn it into knowledge utilised to address social problems (WDR, 1998). As such, developing expertise to utilise big data can equip countries for the twenty-first century challenges of "doing more with less".

In light of this development, big data provide opportunities to bridge the knowledge gap. For governments, extracting value from big data can help in seeking to address public concerns and social issues. Skills of employees must be continuously upgraded to maintain a position. One recent UK government report noted that it is vital that a country's data infrastructure is upgraded to become more "agile, resilient and responsive to changing trends" (DBIS, 2013, p. 5).

Cell II: data specialists (data poor, information rich)

Cell II is where the government collects a limited amount of data, but engages in high-level data analysis to inform public policy. Governments in this category recognise the benefits of data analysis and utilise existing expertise and resources to process the data to inform public policy (see EIU, 2011). An explanation for this could be that the government has sought to counterbalance ethical concerns by collecting very limited data deemed necessary to cater for citizens and inform policy. This can also be attributed to a shift within governments away from large data collection towards high-quality analysis of limited data. Countries with superior expertise in data analytics can turn big data into information, thereby creating and disseminating new knowledge (WDR, 1998; DBIS, 2013). It has been suggested that countries that are able to build capabilities to exploit and use big data intelligently can transform their public service provisions as well as rejuvenate research and development activities (DBIS, 2013). Governments that are able to harness and utilise big data can position themselves for the future by improving quality of decisions.

Cell III: data wasters (data rich, information poor)

Cell III is where governments collect large data but, largely due to government bureaucracy and poor communications within branches of government, the data remains underexplored, underutilised or even unused. In the twenty-first century, it is still not uncommon to observe boxes of notes, floppy disks, cabinets bursting with data sheets and multiple regional data sheets often unprocessed and not connected to any central database (see also WDR, 1998). These are symptoms of governments well-endowed with data, but little integration of data and analysis to inform public

policy. This has been referred to as the "data-rich but information-poor" syndrome (see Ward *et al.*, 1986).

Historically, many governments in the developing world collected and stored data at local centres or regions which were not meaningfully linked together. Retrieving data from local government agencies for analysis can also be costly as well as time consuming (Brennan *et al.*, 2014). A key tenet is that poor communication between branches of government often leaves data unused and underutilised. Often data are stored in silos and disconnected from government departments. As such over time, a lack of integration of multiple data means that the ability to extract maximum value from data was limited. This can also be attributed to the poor or lack of skilled personnel to perform data analysis.

Another explanation for this cell is that the ability to amass, analyse and disseminate such vast amounts of big data are often beyond the capacity of many developing countries (WDR, 1998). Governments may have recognised the importance of big data, but lack of technical expertise hampers data analysis activities leading to poor information and knowledge to feed into public policy formulation.

In spite of recognition of the importance of data collection, lack of quality data analysis can perpetuate the status quo where data have not been efficiently utilised to provide required information (Ward *et al.*, 1986). In this cell, countries' ability to link data collection to inform public policy is limited. Indeed, countries in transition tend to have greater government intervention and bureaucracy which influence the degree of data collection and analysis (see Cavusgil *et al.*, 2012). Governments may have failed to strike the right balance between ethical concerns and need for big data. In so doing, large volumes of data may not be required to perform the functions of a government.

Governments in much of the developing world, lacking in technology expertise and data analysis, are more likely to be found in this category. Although nation state or organisations might collect vast quantities of data, "it is often neither directly useful nor immediately accessible to provide supporting information for many of its clinical and managerial functions" (Forte, 1994, p. 148). Consequently, inability to identify and tap into the data can deprive a country of the potential benefits of big data.

To move from this cell, there is a need for joined-up thinking with governments and electronic recoding and dissemination of data to make it available to researchers. Effective communication between branches of government and frontline employees is essential in overcoming the constraints here. In this direction, recent technological advancement in data recording and storage provides enticing opportunities for countries to seek to update their expertise.

Cell IV: data dropouts (data poor, information poor)

Cell IV is classified as data dropout, where a country engages in low-level data collection and data analysis. The ability to generate new knowledge and understanding from data is curtailed or limited. Consequently, the link between large-scale data collection and public policy formulation appears weak. A possible explanation for this is lack of technological expertise and manpower to engage in data collection and analysis. Governments in this category are characterised by lack of effective strategy to collect, analyse and make sense of data. In many instances, the countries' capabilities and skills of employees have not kept abreast with the evolution of big data, leading to poor understanding of how such data can be utilised to inform public policy. The technological advancement has meant that new data are generated at an accelerated pace which can be captured to improve decision making.

Nonetheless, the ability to analyse and act on big data depends on factors such as skilled workforce, and computing and mathematical skills (DBIS, 2013; Krumholz, 2014). Historical criteria and approaches to data collection and analysis may have become obsolete in the era of big data. For countries in this cell, skills upgrading and retooling of government employees' skills are essential for future success. This also means government needs to invest in training and development of workers to upgrade their skills.

As Yiu (2012, p. 23) observed, for governments "to fully realise the benefits of bringing diverse data sets together, individual data owners must be incentivised to manage the quality of data collection and data capture across the board – and not just for the indicators they (or their managers) deem important". Based on the above discussions on the framework, we offer the following propositions:

P1. When government is characterised by a lack of effective communication between departments, there is more likely to be a poor utilisation of big data.

P2. When government lacks the capacity and capabilities to analyse data on a national scale, this is more likely to hamper its ability to harness big data.

Discussion and conclusions

This paper extends research on big data by examining the different capacities and capabilities of governments in harnessing and utilising big data. The study also sheds light on the relationship between information and data. The conceptualisation led to the development of a 2×2 framework which classifies different capacities and positions of governments relative to data collection and data analysis. The classifications are "data dropouts", "data wasters", "data specialists" and "strategic data users". The framework also highlights the possibilities that any country can upgrade or downgrade its capabilities and expertise, which then affect its position and ability to extract maximum value from big data.

This paper adds to the body of work that suggests such data can be utilised to improve quality of decisions within governments and organisations (Brennan *et al.*, 2014; DBIS, 2013; Amankwah-Amoah, 2015b). Although scholars have explored big data in information systems (Chen *et al.*, 2012; Goes, 2014), there has been a lack of systematic examination of how big data can be harnessed to help governments make informed public policy decisions.

By integrating the literature, the paper articulates conditions for countries to upgrade their capabilities to become "data rich and information rich". The analysis also indicates that there are a number of benefits for governments seeking to become strategic data users. An important limitation of the 2×2 matrix is that it overlooks some of the underlying complex issues that underpin government structure, regulations and cultural issues that shape a country's ability to consolidate data from various sources, communicate data, collect data and utilise data.

Aside from this limitation, two promising lines of research could be pursued to enhance our understanding. First, there is a need for a much broader and extension of the framework developed here. The study lays a foundation towards more robust and sophisticated conceptualisations of different capacities and positions of governments in harnessing and utilising big data. In addition, more work is needed to incorporate ethical issues related to big data and big government which have been largely overlooked here. Big government with big data can exert an enormous influence over the activities of citizens.

One of the potential benefits of this conceptualisation is that it helps to further advance the established view that big data does not necessarily translate into big knowledge and understanding (McAfee and Brynjolfsson, 2012; Snijders *et al.*, 2012). Data analysis presents an indispensable starting point towards knowledge generation and diffusion across government agencies. It is hoped that this study serves as a catalyst for further more scholarly works on the relationship between governments and big data.

References

Amankwah-Amoah, J. (2014a), "Old habits die hard: a tale of two failed companies and an unwanted inheritance", *Journal of Business Research*, Vol. 67 No. 9, pp. 1894-1903.

Amankwah-Amoah, J. (2014b), "A unified framework of explanations for strategic persistence in the wake of others' failures", *Journal of Strategy and Management*, Vol. 7 No. 4, pp. 422-444.

Amankwah-Amoah, J. (2015a), "Governments, airlines and employees: an evolving relationship from 1940 to 2010", *Management & Organizational History*, Vol. 10 No. 1, pp. 1-20.

Amankwah-Amoah, J. (2015b), "A unified framework for incorporating decision-making into explanations of business failure", *Industrial Management and Data System*, Vol. 115 No. 7, pp. 1341-1357.

Amankwah-Amoah, J. and Debrah, Y.A. (2014), "Air Afrique: the demise of a continental icon", *Business History*, Vol. 56 No. 4, pp. 517-546.

Boyd, D. and Crawford, K. (2012), "Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society*, Vol. 15 No. 5, pp. 662-679.

Brennan, N., Oelschlaeger, A., Cox, C. and Tavenner, M. (2014), "Leveraging the big-data revolution: CMS is expanding capabilities to spur health system transformation", *Health Affairs*, Vol. 33 No. 7, pp. 1195-1202.

Cavusgil, S.T., Knight, G.A. and Riesenberger, J.R. (2012), *International Business: Strategy, Management, and the New Realities*, 2nd ed., Pearson Prentice Hall, Upper Saddle River, NJ.

Chen, H., Chiang, R.H. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.

(The) Economist (2010), *Data, Data Everywhere: A Special Report on Managing Information*, Economist Newspaper, London.

Economist Intelligence Unit (EIU) (2011), *Big Data: Harnessing a Game-Changing Asset*, Economist Intelligence Unit, London.

Economist Intelligence Unit (EIU) (2012), *Big Data: Lessons from the Leaders*, Economist Intelligence Unit, London.

Forte, P. (1994), "Data rich, information poor: data, information and decision support in the NHS", *European Journal of Information Systems*, Vol. 3 No. 2, pp. 148-154.

George, G., Haas, M.R. and Pentland, A. (2014), "Big data and management", *Academy of Management Journal*, Vol. 57 No. 2, pp. 321-326.

Ghani, K.R., Zheng, K., Wei, J.T. and Friedman, C.P. (2014), "Harnessing big data for health care and research: are urologists ready?", *European Urology*, Vol. 66 No. 6, pp. 975-977.

Goes, P.B. (2014), "Big data and IS research", *MIS Quarterly*, Vol. 38 No. 3, pp. iii-viii.

Kim, G.H., Trimi, S. and Chung, J.H. (2014), "Big-data applications in the government sector", *Communications of the ACM*, Vol. 57 No. 3, pp. 78-85.

Krumholz, H.M. (2014), "Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system", *Health Affairs*, Vol. 33 No. 7, pp. 1163-1170.

McAfee, A. and Brynjolfsson, E. (2012), "Big data: the management revolution", *Harvard Business Review*, Vol. 90 No. 10, pp. 60-68.

Mayer-Schönberger, V. and Cukier, K. (2013), *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, John Murray Publishers, London.

Milinovich, G.J., Magalhães, R.J.S. and Hu, W. (2015), "Role of big data in the early detection of Ebola and other emerging infectious diseases", *The Lancet Global Health*, Vol. 3 No. 1, pp. e20-e21.

Millward, R. (2005), *Private and Public Enterprise in Europe: Energy, Telecommunications and Transport, 1830-1990*, Cambridge University Press, Cambridge.

Snijders, C., Matzat, U. and Reips, U.D. (2012), "Big data: big gaps of knowledge in the field of internet science", *International Journal of Internet Science*, Vol. 7 No. 1, pp. 1-5.

Tapscott, D., Lowy, A. and Ticoll, D. (2000), *Digital Capital: Harnessing the Power of Business Webs*, Harvard Business Press, Boston, MA.

UK Department for Business, Innovation and Skills (DBIS) (2013), *Seizing the Data Opportunity: A Strategy for UK Data Capability*, Department for Business, Innovation and Skills, London.

Ward, R.C., Loftis, J.C. and McBride, G.B. (1986), "The 'data-rich but information-poor' syndrome in water quality monitoring", *Environmental Management*, Vol. 10 No. 3, pp. 291-297.

World Development Report (WDR) (1998), *Knowledge for Development*, World Bank, Washington, DC.

Welch, C. (2000), "The archaeology of business networks: the use of archival records in case study research", *Journal of Strategic Marketing*, Vol. 8 No. 2, pp. 197-208.

Woerner, S.L. and Wixom, B.H. (2015), "Big data: extending the business strategy toolbox", *Journal of Information Technology*, Vol. 30 No. 1, pp. 60-62.

Yiu, C. (2012), *The Big Data Opportunity: Making Government Faster, Smarter and More Personal*, Policy Exchange, London.

Further reading

BBC (2013), "Big data: predicting and preventing rugby injuries", available at: www.bbc.co.uk/news/health-24836161 (accessed 24 March 2015).

Bhimani, A. (2015), "Exploring big data's strategic consequences", *Journal of Information Technology*, Vol. 30 No. 1, pp. 66-69.

Markus, M.L. (2015), "New games, new rules, new scoreboards: the potential consequences of big data", *Journal of Information Technology*, Vol. 30 No. 1, pp. 58-59.

Newman, A.A. (2011), "Using Google's data to reach consumers", *New York Times*, Vol. 161 No. 55628, 3pp.

The White House (2014), *Big Data and Privacy Working Group Review*, Office of the Press Secretary, Press release Washington, DC.

World Development Report (WDR) (1997), *The State in a Changing World*, World Bank, Washington, DC.

Corresponding author
Dr Joseph Amankwah-Amoah can be contacted at: joeaamoah@gmail.com

This article has been cited by:

1. Joseph Amankwah-Amoah. 2016. Emerging economies, emerging challenges: Mobilising and capturing value from big data. *Technological Forecasting and Social Change* **110**, 167-174. [CrossRef]
2. Youngseok Choi, Habin Lee, Zahir Irani. 2016. Big data-driven fuzzy cognitive map for prioritising IT service procurement in the public sector. *Annals of Operations Research* . [CrossRef]