Government Information Quarterly xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments

Akemi Takeoka Chatfield [a,*], Christopher G. Reddick [b]

[a] School of Computing and Information Technology, Faculty of Engineering and Information Sciences, University of Wollongong, Australia
[b] Department of Public Administration, College of Public Policy, The University of Texas at San Antonio, United States

ARTICLE INFO

Article history:
Received 26 October 2016
Received in revised form 15 January 2017
Accepted 10 February 2017
Available online xxxx

Keywords:
Longitudinal cross-sector analysis
Open data portal
Service capability
Open data
Big data
Data analytics
Open innovation
Open data policy intensity
Citizens/portal users

ABSTRACT

While open government partnerships and open government data initiatives around the world have proliferated in practice, empirical research is required to better understand open data policy and open data portal capability which would spur meaningful citizen engagement towards co-production of open services innovation through open data reuse. Specifically, relatively little has been empirically investigated about open data portal as supply-side service capabilities at the local government level. In this longitudinal research on twenty open data portals in Australia's largest cities, cross-sector analysis results find large variation in open data portal service capabilities, which are measured by open data policy intensity, open data provision, data format variety, and entrepreneurial data services, including analytics tools, data modeling, and hackathon idea competitions. Longitudinal cross-sector analysis results also find the important roles played by open data policy and dedicated open data portal investment as predictors of open data portal service capability improvements over time.

1. Introduction

Open government data (OGD) has been seen as new mechanisms for achieving government transparency, civic engagement, and new forms of collaboration for open innovation which were primary goals of the Open Government Directive of the former Obama administration (U.S. Executive Office, 2009). Many other governments throughout the world have also implemented OGD initiatives (Janssen, Charalabidis, & Zuiderwijk, 2012; Mergel & Desouza, 2013; Kassen, 2013; Zuiderwijk & Janssen, 2014; Conradie & Choenni, 2014; Meijer, Conradie, & Choenni, 2014; Luna-Reyes, Bertot, & Mellouli, 2014; Veljković, Bogdanović-Dinić, & Stoimenov, 2014; Jetzek, Avital, & Björn-Andersen, 2014; Open Government Partnership, 2016). The concept of OGD underscores the recognition of OGD as valuable tangible or intangible resources at the government's disposal (Open Government Working Group, 2007; Alanazi & Chatfield, 2012).

The transformative potential of big and open data is notable for enhancing e-government services, openness, government transparency, citizen engagement, and the interaction between governments, citizens, and businesses (Bertot, Gorham, Jaeger, Sarin, & Choi, 2014). Moreover, it is widely held that reuse of OGD has the potential to generate open innovation (Zuiderwijk, Janssen, & Davis, 2014; Susha, Grönlund, & Janssen, 2015) and both economic and social values (Jetzek et al., 2014) in the big data and linked open data ecosystems (Mayer-Schönberger & Cukier, 2014; Chen, Mao, & Liu, 2014). In the short period, open government, open data, and open data policy research streams have gained traction (Janssen et al., 2012; Lee & Kwak, 2012; Kassen, 2013; Evans & Campos, 2013; Luna-Reyes et al., 2014; Zuiderwijk & Janssen, 2014; Veljković et al., 2014; Bates, 2014; Martin, 2014). The empowering potential of OGD realized at the local government level may provide a useful platform for promoting proactive citizen engagement (Kassen, 2013). Rather than an emergent view of open data portal as a "public e-service" (Lněnička, 2015, p. 589), however, there exists a simplistic view of open data which equates the government's open data provision with citizen engagement in reusing open data to create benefits (Janssen et al., 2012). Hence, the promotion of meaningful citizen engagement is one of the key challenges facing open government initiatives (Ganapati & Reddick, 2012; Luna-Reyes & Ae Chun, 2012; Evans & Campos, 2013).

Linders (2013) argues that while open data has made substantial contributions towards realizing a more integrated vision of international aid delivery, much of the potential of open data remains unexplored. Furthermore, Kassen (2013, p. 509) also argues that "in practice it is not yet clear how the potential of the open data concept can be realized at the local level as there has been no analysis of current projects thus far." To date, however, "a comprehensive analysis of the capabilities

* Corresponding author.
E-mail addresses: akemi@uow.edu.au (A.T. Chatfield), chris.reddick@utsa.edu (C.G. Reddick).

http://dx.doi.org/10.1016/j.giq.2017.02.004
0740-624X/

and potential of these initiatives is currently missing from the recent research literature" (Petychakis, Vasileiou, Georgis, Mouzakitis, & Psarras, 2014, P. 34). There remains "the lack of a clear way to enable empirical analysis and quantitative measurements of OGD initiatives (Carrasco & Sobrepere, 2015, p. 633). Moreover, "little has been done to analyze and prove the impact and accrued value of these OGD initiatives." (Ubaldi, 2013, p. 1) Finally, empirical research on open data portals at the sub-national level is still lacking (Thorsby, Stowers, Wolslegal, & Tunbuam, 2016).

To address the problem of these observed knowledge gaps, we raise a central research question in this paper:

Do open data portal service capabilities differ in terms of open data provision, data format variety, open data policy intensity, and entrepreneurial data services, including analytics tools, hackathon competitions, and data modeling?

In addressing the research question, we draw on prior OGD research to empirically analyze services and characteristics of OGD portals. In this paper, we define capability as the sophistication of open data service in terms of the key characteristics of the portal listed above in our research question. Prior research on open data portals identified the complexities of the issues involved in the implementation and operation of dynamically changing open data portals, including inadequate rewards for sharing data (Reichman, Jones, & Schildhauser, 2011), inadequate open data policy frameworks for big data (Bertot et al., 2014), data quality controls (Vetrò et al., 2016), data category standard (Thorsby et al., 2016), stimulating civic apps development (Lee, Almirall, & Wareham, 2015), and local challenges (Kassen, 2013; Conradie & Choenni, 2014). However, prior research including Thorsby et al. (2016) limited an analysis of open data portals to only one specific point of time. Since services and characteristics of open data portals are dynamically changing over time, we adopted a longitudinal cross-sector analysis methodology in evaluating twenty Australian open data portals at the local government level over two different time periods. In this paper, we aim to develop a better understanding of supply-side open data portal service capabilities which we view as essential for stimulating demand-side responses: citizen engagement and citizen co-production of open innovation.

The structure of the remaining paper is as follows. The second section presents theoretical foundation that aims to provide us a roadmap for conducting a longitudinal cross-sector analysis of services and characteristics of the operational open data portals. The third section describes the longitudinal research methodology adopted to answer the central research question. The fourth section presents our key findings. The fifth section presents our discussion and the sixth section presents our conclusions and policy recommendations.

2. Theoretical foundation

In order to answer our research question on the supply-side open data portal performance dynamics, we have drawn on prior research on open data portals (Kassen, 2013; Attard, Orlandi, Scerri, & Auer, 2015; Thorsby et al., 2016), OGD (Bertot et al., 2014), data analytics tools (Pauwels et al., 2009), and open data policy (Zuiderwijk & Janssen, 2014) to identify salient services and characteristics of open data portals developed and implemented by local governments for proactive publication of open datasets owned by their various departments.

2.1. Open data portal

An open data portal "allows users to publish, manage and consume data in machine-readable formats, interlink their data with data published elsewhere on the Web, publish applications built on top of the data, and interact with other users" (Kostovski, Jovanovik, & Trajanov, 2012, p. 2). In contrast to research streams on OGD opportunities and challenges and open data policy frameworks that have gained traction, a comprehensive analysis of the capabilities and potential of these open data portals is currently lacking in the research literature (Petychakis et al., 2014). Moreover, open data portal development at the local government level remains at an early stage even in the developed nations such as the U.S. (Thorsby et al., 2016).

Despite the early stage of development, empirical research on open data portal analysis is emerging. Table 1 shows a summary of the existing empirical research on open data portal analysis in terms of the level of government, measures of open data portal services and characteristics, and whether or not open data policy was evaluated.

Despite the early stage of development, Table 1 shows the recent research interests in understanding open data portal capabilities through an analysis of big data such as sensor data and geospatial data published by local governments (Bui, 2015; Okamoto, 2016; Oliveira & Moreno, 2016). Table 1 also shows that the recent open data portal analysis research, except Thorsby et al. (2016), fail to examine the presence/absence of open data policy as part of the open data portal characteristics.

An analysis of features and content of 36 open data portals in American cities find that the portals are in a very early stage of development and the provision of data visualization and data analytics tools is lacking across the cities (Thorsby et al., 2016). Regression analysis results find that population size is the best predictor of the number of datasets published, explaining 79.8% of the variation in the number of datasets. In contrast, other independent variables such as the age of the portal, the type of government, the degree of civic innovation and the level of education in the city were not significant in predicting the number of datasets published on the data portal.

A study of seven national open data portals as tools for transparency and accountability in U.S., U.K., Canada, France, Australia, New Zealand, and Singapore found that overall the portals functioned as "simple data repositories", which failed to facilitate ordinary citizens to reuse open data without the provision of high quality datasets and a complete listing of metadata fields (Lourenço, 2016, p. 331). Furthermore, the practice of releasing open data faced challenges of task complexity and data quality in a case analysis of U.S. Department of Defense contracting data (Whitmore, 2014).

2.2. Open government data

OGD refers to as "data that are freely available to everyone to use and republish without restrictions from copyright, patents or other mechanisms of control" (Jaakkola, Mäkinen, & Eteläaho, 2014, p.26). There are three principles underlying this definition: (1) availability and access; (2) re-use and redistribution; and (3) universal participation (Maccani, Donnellan, & Helfert, 2015).

The five-star deployment scheme for open data was developed by Tim Berners-Lee (Five-Star Open Data, 2015), which is useful to classify the type of data formats among the datasets published on the open data portals. One-star indicates that while open data is machine readable, the data in a table or a chart is "locked-up in a document" such as PDF and is difficult for portal users to reuse the data, and hence providing them with lowest-level value. Two-star indicates structured data is published in a proprietary format such as Excel's XLS that would require proprietary application software for data reuse, aggregation, computation and visualization. Three-star indicates that open data is published in a non-proprietary open format such as CSV as well as of XLSX, enabling a greater number of portal users to reuse and process data with tabular structure (XLS/CSV) and tree structure (XML/JSON). Four-star indicates that unstructured data is published in a RDF graph structure and URLs are assigned to open data for linking with other data. Five-star indicates linked open data through which portal users can uncover other related data and explore the data schema directly through APIs, and hence providing them with highest added value of OGD towards apps development and open services innovation.

Sieber and Johnson (2015) proposed four models of open data provision: data over the wall, code exchange, civic issue tracker, and

Please cite this article as: Chatfield, A.T., & Reddick, C.G., A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments, *Government Information Quarterly* (2017), http://dx.doi.org/10.1016/j.giq.2017.02.004