



Contents lists available at ScienceDirect

Government Information Quarterly

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

Understanding the content and features of open data portals in American cities

Jeffrey Thorsby, Genie N.L. Stowers*, Kristen Wolslegel, Ellie Tumbuan

Public Administration Program, School of Public Affairs and Civic Engagement, Downtown Campus, Suite 679, San Francisco State University, 835 Market Street, San Francisco, CA 94103, United States

ARTICLE INFO

Article history:

Received 24 June 2015

Received in revised form 9 July 2016

Accepted 14 July 2016

Available online xxxxx

Keywords:

Open data

Open data portals

Open government

Civic technology

Data portals

ABSTRACT

In this paper, we present the results of research on features and content of open data portals in American cities. Five scales are developed to categorize and describe these portals: the Open Data Portal Index (ODPI), Data Content Index (DCI), a compilation of the two (Overall Index), the Number of Datasets and Number of Datasets per 100,000. Regression models explaining variation between cities on these scales indicate city population as an important influence, along with participation in a regional consortium. More variation could be explained in the number of datasets model (79.8%) than in any other model. Overall, results indicate portals are in a very early stage of development and need a great deal of work to improve user help and analysis features as well as inclusion of features to help citizens understand the data, such as more charting and analysis.

© 2016 Published by Elsevier Inc.

While descriptive research has been conducted on open data and open data portals at the national level, and in other countries at the national and sub-national levels, there is still a need to conduct basic empirical research upon American urban open data portals. In this article, we present not just descriptive research on these portals but also much needed results of inferential analysis of the characteristics of the first open data portals in American cities. The article begins with a description of open data, open data portals and a brief history of their beginnings. Methods for the study are then discussed and characteristics of the first portals are described, based upon examination of the first 37 urban open data sites to be posted on the Cities site of data.gov. Indices based on these features are described and regression models posited to explain the variation among cities in these models. Finally, the results of these models are described.

1. Open data and data portals

The Internet has permanently altered the process by which information is obtained on a daily basis, particularly within the realms of news consumption and business commerce. The term *Web 2.0*, coined by Tim O'Reilly (2005) to describe the second wave of innovation in information and communication technologies (ICTs), utilizes advanced participatory features such as podcasts, widgets, and Really Simple Syndication (RSS). More recently, the Internet has begun to influence how societies use ICTs within their governments. Sandoval-Almazan

and Gil-Garcia (2012) point out that the definition of electronic government (e-government) has been and continues to be debated by scholars. On one end, it is narrowly defined, with its purpose being simply to enhance government services. On the other end, it is broadly defined and includes various degrees of participatory engagement aimed at increasing democratic systems and empowering citizens.

With the emergence of Web 2.0, e-government has been presented with the opportunity to take an active role in further reshaping its definition. Taewoo Nam (2012) explains that e-government now includes Government 2.0 and Open Government as the means and the ends of current e-government rationale. Specifically, "the transition of government [breaks down] into new modes in terms of goals and tools. The new aspects of e-government are not only for [the] government, but also for the public as customers and users" (p. 347).

2. The principles of open data

In 2009, a group of activists calling themselves the Open Government Working Group laid out eight principles for the use of public data. They proposed data was open when it is: 1) complete; 2) primary; 3) timely; 4) accessible; 5) machine-processable; 6) non-discriminatory; 7) non-proprietary; and, 8) license-free (Dawes, 2010). With these principles, they believed data would be available to all without the need for purchase or special software, its source and purpose would be transparent, and data could be used by anyone who wanted to access it, thereby enhancing participation and collaboration around the topic of the data.

* Corresponding author.

E-mail address: gstowers@sfsu.edu (G.N.L. Stowers).

Further movement towards open data was the shift of U.S. federal policy towards e-government as “open government”, underscored in President Obama’s Memorandum on Transparency and Open Government in 2009. This policy document defined three foundational principles: transparency, participation and collaboration (Ganapati and Reddick, 2012, p. 115; McDermott, 2010; Jaeger and Bertot, 2010).¹

In 2010, Tim Berners-Lee (the founder of hypertext linking) extended our understanding by laying out the levels of open data with his Linked Open Data 5 Star plan (Berners-Lee, 2010). This scheme is found in Table 1. One star data is available to all, two stars mean the data is machine-readable, three stars means it is machine-readable and in a non-proprietary format. With four stars, the data has the properties of the three previous levels plus is based upon standards (RDF, or Resource Description Framework) which allow precise pointing to the data online. Five stars, or the most accessible open data level, includes everything else plus the ability to link one’s data to another’s data.

Later scholars and activists further refined open data definitions with three related principles: *availability and access*, *reuse and redistribution* and *universal participation* (Gurstein, 2013). Gurstein argues that data should be made freely available for no more than its reproduction cost, in its full form, and inconvenient and modifiable machine-readable formats (*availability and access*). Providing data more freely might mean using the non-proprietary CSV format, which is modifiable and widely used today on open data portals. Further, data need to be made available under terms that allow for their reuse and redistribution, including the ability to combine data with other datasets (*reuse and redistribution*). A CSV formatted dataset would enable users to download the dataset, combine with other CSV datasets and redistribute to other users interested in the same issue. Lastly, data must also be made available to the public in an equitable fashion (*universal participation*), such as a simpler format that would allow users without advanced technical skills, or more advanced software, to use the data. Machado and De Oliveira (2011) further defined open data as “the publication of data in open raw formats and ways that make it accessible and readily available to all and allow reuse, such as the creation of data mashups and applications” (p. 449).

The three laws of open government data (herein referenced as ‘open data’ for this paper) summarize what format it should take, and why:

1. If it [(data)] can not be spidered or indexed, it does not exist.
2. If it [(data)] is not available in open machine readable format, it can not engage.
3. If a legal framework does not allow it [(data)] to be repurposed, it does not empower (Eaves, 2009).

Kassen (2013) noted that the value of open data developed through releasing government data to the public in machine-readable formats because presumably both government transparency and civic engagement will increase. Such reasoning is consistent with Tolbert and Mossberger’s (2006) findings that e-government initiatives have a positive impact on citizen attitudes of trust in their local government. Even more specifically, the positive effects of open data can be broadened to include economic innovation, increased accountability, increased government agency efficiency, and social and economic innovation (Kitchin, 2013). In just two cities (New York and Chicago), numerous instances have been cited of ways open data has benefitted city government or citizens themselves. New York officials point to the analysis of open data about building inspections to help improve predictions of fires; Chicago points to their use of crime data and the development of helpful applications from their open data (Goldstein and Dyson, 2013).

Despite the potential of positive changes towards accountability and transparency that open data could bring, concerns are increasingly

¹ McDermott (2010) explains that the Open Government Initiative is legal based on pre-existing laws including, but not limited to, The Paperwork Reduction Act of 1980, The E-Government Act of 2001 and 2002, The Freedom of Information Act and its amendments and the Open Government Act, 2007 (p. 404–411).

Table 1

Tim Berners-Lee linked open data 5 star scheme.

Source: Berners-Lee, Tim. (2010). Is your linked open data 5 star? Retrieved on October 16, 2015, from <http://www.w3.org/DesignIssues/LinkedData.html>.

Stars	Interpretation
*	“Available on the web (whatever format) but with an open license, to be Open Data
**	“Available as machine-readable structured data (e.g., excel instead of image scan of a table)
***	“as (2) plus non-proprietary format (e.g. CSV instead of excel)
****	“All of the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
*****	“All the above, plus: Link your data to other people’s data to provide context”

being raised about the intended and/or unintended consequences of open data. For example, Kitchin (2013) raises concerns over neoliberalization and marketization of public services as open data is increasingly used primarily to build applications by private sector firms, further exacerbating inequality. While activists hope that open data can further citizen engagement with their governments, little evidence so far exists that this is happening.

Critics have raised other concerns, such as the important distinction between open data and the Right to Information Movement (RTI movement), “which promotes access to government information as a fundamental right” (Janssen, 2012, p. 3). Specifically, Janssen contends that if governments predominantly rely on economic innovation as a vehicle for transparency and accountability, then “open data risks the illusion of transparency and accountability, while in reality [it could be] causing information inequality and disempowerment of citizens” (p.12). The fact that there exists a digital divide, where access to information technology is not equitably distributed across populations, gives credence to this argument. Bates (2012) similarly concludes that a real danger of open data initiatives is that they would become framed in a manner that “champions the superiority of markets over social provision...potentially becoming, little more than corporate subsidy” (p. 10).

Such concerns raise a potential conflict between the goals of accountability and transparency and those of economic stimulation. Further, in this era of information scandals such as the case of Edward Snowden, who leaked information about a National Security Agency program that was spying on citizens within the United States (Risen, 2013), urgent concerns about government and transparency are apparent. Janssen (2012) argues that the type of data released through open data platforms has a direct effect on the outcome of accountability. Geographic data, postcodes, transportation data, corporate data and business data will stimulate economic innovation; whereas, budgetary data, spending expenditures, procedural data, and legal data will stimulate accountability efforts (p. 11).

3. Open data evolution

In 2007, a group of open software activists met in Sebastopol, California and outlined the basic concepts of open data, defining public data as a public good that should be available to all (Chignard, 2013). President Obama, in his 2009 executive order, codified their ideas, a groundswell of support emerged around the Open Government concept, and the federal government created data.gov. Collectively called the open data movement, it has sprung up throughout the majority of states, in many cities throughout the country (Townsend, 2013) and now, across the world.

The CIOs of Chicago, Los Angeles, Boston, New York, Seattle, Washington, D.C. and San Francisco, working informally as the Group of Seven in 2009, focused on the need for a common open interface for

Download English Version:

<https://daneshyari.com/en/article/5110664>

Download Persian Version:

<https://daneshyari.com/article/5110664>

[Daneshyari.com](https://daneshyari.com)