

Crowdsourcing open data policies: Measuring impacts and improving outcomes

Gwen Shaffer
California State University Long Beach
gwen.shaffer@csulb.edu

Abstract. THIS RESEARCH EXPLORES BOTH THEORETICAL AND PRACTICAL ASPECTS OF DEVELOPING AND IMPLEMENTING OPEN DATA POLICIES PRODUCED THROUGH COLLABORATIVE GOVERNANCE. IT EXPLORES HOW CITY OFFICIALS CAN MAXIMIZE TIME AND RESOURCES INVESTED IN SEEKING PUBLIC INPUT THAT WILL RESULT IN BETTER OPEN DATA POLICIES. AS A KEY CASE STUDY, THE AUTHOR EXAMINES EFFORTS TO DRAFT AN OPEN DATA POLICY IN THE CITY OF LONG BEACH, CALIF., WHERE OFFICIALS HOSTED A SERIES OF OPEN DATA FORUMS AND DISTRIBUTED A SURVEY DURING FALL 2015.

1 Introduction

During the past few years, local governments have undertaken a host of efforts to crowdsource open data policies. This position paper calls for an exploration into both the theoretical and practical aspects of developing and implementing open data policies produced through collaborative governance. Since the 1950s, agencies at all levels of government have hosted citizen participation initiatives (Day, 1997). These practices are widely believed to boost public trust of government, as well as foster a stronger sense of community. Simply opening up the policymaking process can serve as a transformative tool for social change (Stivers, 1990; Oldfield, 1990; Nelson & Wright, 1995).

However, this paper highlights the distinct characteristics associated with participatory governance within the context of open data policies. It raises questions regarding when, and how, open data policies should rely on public participation. First, government officials typically initiate collaboration in an attempt to reach consensus in a controversy (Ansell & Gash, 2007) like raising property taxes or closing a library. By contrast, open data stakeholders are not attempting to solve a problem, *per se*. And while stakeholders who help shape open data policies have diverse agendas—running the gamut from an interest in developing mobile applications, to highlighting disparities in education spending—these actors are not adversaries. Finally, this paper questions the effectiveness of local government attempts to crowdsource open data policies through digital platforms, in light of research finding collaborative governance requires face-to-face interactions with the public (Ansell & Gash, 1997).

As of December 2015, about 55 U.S. cities had adopted open data policies, including Los Angeles, New York City, Chicago, Philadelphia, Boston, and San Francisco (Sunlight Foundation, 2015). Open data policies provide the public with guidelines addressing which data should be public and how to make that data public (i.e. standardized formatting, redacting personal information). Because it involves political processes, crafting an open data policy is more complex than publishing government data on a website (Shaw, 2015). Without a codified policy, a newly installed mayor can deem certain data sets sensitive—and remove them from an open data portal, or department staff can decide to stop updating particular data sets. Most significantly, however, open data policies explain the rationale for making government data available to the public. These formal guidelines justify stakeholders' rights to request access to data (Shaw, 2015).

2 Discussion

In the majority of policy contexts, collaboration involves soliciting input from residents who are most impacted by local problems, and who possess unique knowledge about these situations. For example, more than 300 stakeholders representing agriculture, forestry and land development attended meetings and provided feedback when California regulators updated the state's strategic water plan in 2013 (Beutler, 2014). In this scenario, the people most likely to be negatively impacted by water restrictions and rate hikes weighed in on the plan. The dynamics are different, however, in the context of shaping open data policies. For example, the city of Long Beach, Calif., hosted a series of open data forums during Fall 2015. About 70 attendees completed a survey inquiring about their interests and concerns surrounding open data. These respondents reported a range of occupations: web developer, college student, journalist, analyst, retiree, and community activist, among others (Long Beach Technology & Innovation

Commission, 2015). Nearly 60 percent of respondents said they had never downloaded a “local, state or federal government open data set.” While participants clearly had an interest in open data, they generally lacked first-hand experience engaging with open data. The city of Long Beach’s collaborative approach has multiple advantages—informing and empowering residents, as well as building trust—but it remains to be seen whether discussions that took place during open data forums will ultimately benefit Long Beach’s open data policy.

The city of Philadelphia (2015) has taken a different approach to participatory governance by creating an Open Data Advisory Group. In addition to staff from various city agencies (i.e. parks, streets, licenses and inspections), the advisory group includes representatives working in mass transit; healthcare; “good” government advocacy; and business. While each of these sectors anticipates using open data to meet their needs, they lack expertise with open data. This raises questions about how significantly the group’s recommendations will influence outcomes for Philadelphia’s open data initiative.

Still, good reasons for participatory efforts exist. Government officials may host forums and meetings in hopes of “demystifying” the term open data open, or in an effort to promote downloads from open data portals. Cities also seek public input as they struggle to prioritize the release of datasets. Therefore, future research should examine, do the time and resources invested in seeking public input into open data policies actually result in *better* policy? Currently, no formal tools or methods of evaluation exist to measure the benefits and outcomes achieved through public participation.

This paper also argues for distinguishing between “collaborative governance” and “crowdsourcing.” Previous research on open data stresses the need for users to provide feedback on the datasets they perceive as most valuable (Ubaldi, 2013; Kassen, 2013). In fact, most open data portals allow the public to suggest the release of specific data and to offer feedback on the portals themselves (NYC Open Data, 2015; City of Chicago, 2015; District of Columbia, 2015; Nashville Open Data, 2015). These opportunities certainly qualify as *crowd production*. However, it is debatable whether they meet the definition of *collaborative governance*, which typically requires face-to-face dialogue between stakeholders. Soliciting online feedback, hosting surveys, and conducting focus groups are not collaborative in the sense that they do not permit two-way flows of communication or multilateral deliberation (Ansall & Gash, 2007).

Even when local governments undertake efforts that are truly collaborative, downsides may emerge. In 2013 community members in Oakland, Calif., helped write a first draft of that city’s open data policy. The draft was shared through Google docs, and members of the public could mark it up. Following this public comment period, the city hosted an open data roundtable. This event included a Google hangout, allowing remote participation. In many respects, the process that took place in Oakland serves as a model for engaging the community in creation

of open data policy (Williams, 2013). But relying too heavily on public participation may inadvertently reinforce the preferences of people already comfortable interacting with government officials (Sunshine Foundation, 2015). Since the average person knows little about open data, meetings may be dominated by tech-savvy citizens or special interests who disproportionately influence decision-making (Irvin & Stansbury, 2004). In a landmark study on collaborative efforts to draft natural resource policy, Smith and McDonough (2001) found inequality in representation. Informants characterized public meetings as “orchestrated” and “loaded” (p. 245). The results of Long Beach’s open data survey back up these concerns. Among open data forum attendees who completed the city’s survey, 94% reported having at least some college education (Long Beach Technology & Innovation Commission, 2015). This raises the possibility of a small elite dominating a participatory process, unless the city modifies its method of soliciting public participation.

3 Conclusion

While some local governments are involving the public in creation of their open data policies, these efforts tend to fade after an open data portal goes online. This position paper argues that—just as public preferences should be incorporated into the development of an open data policy and prioritizing data sets for initial release—the public should remain involved in the ongoing assessment of the portal itself (Sunlight Foundation, 2014). Governments can crowdsource public feedback regarding data quality, quantity, selection, and publishing format. Of course, this feedback is meaningless if officials fail to incorporate it into policy revisions. Therefore, future research should examine whether cities are, in fact, addressing public concerns when open data policies undergo routine review.

Despite the challenges highlighted in this paper, collaboratively drafted open data policies have multiple potential benefits. They can increase public awareness, as well as empower residents. City officials who engage the public are positioned to develop better policy and make more informed decisions during later implementation phases (Irvin & Stansbury, 2004). As an increasing number of local governments undertake open data initiatives, officials must consider the contextual conditions likely to facilitate—or hinder—the desired outcomes of crowdsourcing policies. By conducting research that measures and clarifies the particular ways in which communities engage with open data, scholars can boost the effectiveness of collaborative governance. Therefore, future open data policies must include tools for evaluating public participation and its impacts. Ideally, these assessment methods will enable cities to reach out to underrepresented groups and obtain more diverse perspectives.

4 References

[1] Ansell, C., and Gash, A. (2007). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory*, 18, 543-571.

[2] Beutler, L. (2014). Digital engagement case study: Collaborative policymaking: California water plan update. Retrieved from <http://www.slideshare.net/IAP2NorCal/digital-engagement-casestudy-california-water>.

[3] City of Chicago (2015). Dataset suggestions. Retrieved from <https://data.cityofchicago.org/nominate>.

[4] City of Minneapolis (2015, April 20). City hosts open data forum April 27. Retrieved from <http://www.ci.minneapolis.mn.us/news/WCMS1P-140339>.

[5] City of Philadelphia (2015, June 11). City's Open Data Advisory Group meets. *PhillyInnovates* blog. Retrieved from <http://phillyinnovates.com/2015/06/11/citys-open-data-advisory-group-meets/>.

[6] Day, D. (1997). Citizen participation in the planning process: An essentially contested concept? *Journal of Planning Literature* 11(3), 421-434.

[7] District of Columbia (2015). Request for comments on the District of Columbia's Transparency and Open Government Initiative. Retrieved from <http://dc.gov/page/request-comments-district> columbias-transparency-and-open-government-initiative.

[8] Irvin, R. and Stansbury, J. (2004). Citizen participation in decision-making. Is it worth the effort? *Public Administration Review*, 64(1), 55-65.

[9] Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly*, 30, 508-513.

[10] Long Beach Technology & Innovation Commission, 2015. Preliminary report on open data surveys. Discussed during the December 16, 2015 commission meeting.

[11] Nashville Open Data Portal (2015). Dataset suggestions. Retrieved from <https://data.nashville.gov/nominate>.

[12] Nelson, N. and Wright, S. (1995). *Power and participatory development: Theory and practice*. London: Intermediate Technology Publications.

- [13] NYC Open Data (2015). Dataset suggestions. Retrieved from <https://data.cityofnewyork.us/nominate>.
- [14] Oldfield, A. (1990). *Citizenship and community: Civic republicanism and the modern world*. London: Routledge.
- [15] Shaw, E. (2015). Making the most of open data. In *What counts: Harnessing data for America's communities*, 108-122. San Francisco: The Federal Reserve Bank of San Francisco and the Urban Institute.
- [16] Smith, P. and McDonough, M. (2001). Beyond public participation: Fairness in natural resource decision making. *Society & Natural Resources* (14), 239–249.
- [17] Stivers, C. (1990). The public agency as polls: Active citizenship in the administrative state. *Administration & Society* 22(1), 86-105.
- [18] Sunlight Foundation (2015). Open data policies at work. Retrieved from <https://sunlightfoundation.com/policy/opendatamap/>.
- [19] Sunlight Foundation (2014, March). *Guidelines for open data policies, version 3*.
- [20] Ubaldi, B. (2013). Open government data: Towards empirical analysis of open government data initiatives. OECD Working Papers on Public Governance, No. 22. Retrieved from <http://dx.doi.org/10.1787/5k46bj4f03s7-en>.
- [21] Williams, R. (2013). Oakland's public participation route to open data legislation. Sunlight Foundation. Retrieved from <https://sunlightfoundation.com/blog/2013/10/29/oaklands-public-participation-route-to-open-data-legislation/>.