

# A 'data curation' model for direct citizen editing of government spatial data

**Peter A. Johnson**

Geography and Environmental Management, University of Waterloo, [peter.johnson@uwaterloo.ca](mailto:peter.johnson@uwaterloo.ca)

## **Abstract**

Throughout Canada, governments at all levels have begun to share their geospatial data as part of a move towards increasing openness and transparency. By providing geospatial data online, for no cost, in a variety of accessible formats, and with minimal restrictions on use, governments are opening up opportunities not only for citizen access, but also for citizen participation. The increasing adoption of mobile devices is one trend that may facilitate the engagement of citizens with government through the provision of open geospatial data. This research presents a framework for using government geospatial data as a conduit for citizen participation through mobile device-based editing of official government data. This research advances a model where volunteered data contributions correcting, editing, or expanding government open data are made through a mobile device and integrated into an authoritative database, after curation by a government worker, with feedback to contributor. This approach of 'data curation' is compared to a crowdsourcing model. Constraints to the adoption of volunteered data within government, including aspects of data quality and provenance that may drive government towards a curation model, are considered.

## **Background and Relevance**

The ways that governments can and do provide data is a rapidly emerging area of research, with direct implications for the relationship between government and citizen. Governments have long collected information, including geospatial data, with which to support planning, decision-making, and service provision (Janssen, Charalabidis, & Zuiderwijk, 2012; Tinati, Carr, Halford, & Pope, 2012). Traditionally this data was kept internal to the organization and only made publicly available in a distilled, generalized format, if at all. The widespread availability of public sector data on the Internet represents a shift towards opening and distributing datasets for general public and private sector access and use (Yu & Robinson, 2012). More fundamentally, it represents a transformation over time in the value of government data, from a means to an end in policy deliberations, to an end in itself (Onsrud & Pinto, 1991). Open data is provided as a way to facilitate access to government data, and it is argued that through this provision, increased participation in government functioning, decision-support, and service provision can result.

The provision of open data is grounded, from the government side, on e-government and freedom of information (FOI) initiatives and, from civil society, of increasing pressure for a more open government based around principles of efficiency, transparency, and communication with citizens (Gurstein, 2011; McDermott, 2010). Open data is derived from an underlying ideology of openness, as a way to advance the principles of open government (Ganapati & Reddick, 2012; Meijer, Curtin, & Hillebrandt, 2012). This embedding of open data as a key component of a more broadly open government shows that open data is a concrete step towards both transparency and affording civil society a right to check on government activity and a way

for the public to participate as a collaborator within the processes of service provision, governance, and decision-making.

As mobile devices take an increasingly central role in our daily activities, they become mediators of our interactions with each other, the environment, businesses, and institutions (e Silva & Frith, 2010). With this newly realized ubiquity, there are many resulting opportunities for the creative engagement of citizens within a reciprocal exchange of information with government. The generation of volunteered geographic information (VGI) has been widely explored in a number of contexts (Elwood, Goodchild, & Sui, 2012; Sui & Delyser, 2011), yet questions remain as to how VGI can be integrated into formal governance and decision-making channels (Johnson & Sieber, 2012b).

## **Methods and Data**

To connect VGI as a support for government decision-making, two high-level approaches are possible; crowdsourcing or 'data curation'. This paper aims to provide a comparison of these two approaches within a context of citizen editing of official government base map data. First, the crowdsourcing model follows the template laid out by notable crowdsourced information repositories, such as Wikipedia and Openstreetmap. Under this model, contributions are made directly to the information source (text or spatial features), in real-time. Contributions are then evaluated by a community of users, who may choose individually to re-modify, alter, or even erase the contribution. With the crowdsourcing model, a detailed history of changes (a changelog) is kept, showing how a given document or feature has evolved. Comparably, a 'data curation' model relies on an authoritative 'expert' to vet or otherwise evaluate a given citizen-proposed change before being accepted (or formally committed) to the master database. In a government data context, this expert curator would be most likely a government employee or team with responsibility for data management and updating. This data curation model is positioned as a compromise between the crowdsourced model, and the traditional model of closed government data created with little direct citizen input.

This 'data curation' model contradicts much of the current enthusiasm for the potential of crowdsourced and community-vetted contributions to create high-quality, authoritative knowledge (Brabham, 2009; Dodge & Kitchin, 2013). In the geospatial data/VGI literature there is a significant focus on establishing that data contributed meets expected levels of quality or accuracy (Girres & Touya, 2010; Goetz & Zipf, 2012). Though this thrust of research has substantial value, it is argued that for government use in such critical areas as infrastructure management, health care, services planning, and emergency plan development, that such crowdsourced data faces significant constraints to adoption (Johnson & Sieber, 2012a; 2012b). Most notably, though currency of crowdsourced data may have advantages compared to traditional datasets, the quality, accuracy, provenance, and even legitimacy of crowdsourced data generally remains opaque (Feick & Roche, 2012; Haklay & Ellul, 2011). These are all identified as critical barriers to the adoption of crowdsourced data by government operations. It is as a result of these intractable barriers that an alternate approach to engaging citizens in the creation and sharing of data with government is required.

## **Results**

This paper presents a data curation model, where volunteered data contributions are integrated into an authoritative database with feedback to contributor. This model allows asserted information to be incorporated into an authoritative map product. The data curation model is tested first through a university campus-scale deployment of mobile device feature editing application. Research participants will be tasked with editing and updating a number of campus

geodatabase features using a mobile device feedback application. This application is embedded within a workflow where campus facilities management staff respond to edits, in a facsimile of the citizen-government relationship, with the campus base map serving as a form of open data. From this preliminary stage research, there are several areas of results, including a focus on deploying and testing the mechanics of mobile device application use for geospatial data editing. This work identifies user interface challenges, including manipulating geometries, as constraints to the use of mobile devices in this fashion. The data curation model is assessed, in how it differs from the crowdsourcing model in terms of user motivation to contribute, data quality and completeness, and most critically, how a curated model addresses government adoption constraints.

## Conclusions

This research serves as an empirical example of the use of mobile devices and collaborative base map editing as a realization of open government principles. Though the crowdsourcing model of integrating VGI into government receives much attention, this model is not entirely in line with the nature and regulations of government use and custodianship of critical data. Alternately, a data curation model is proposed as a way to vet citizen contributions before merging into an authoritative database. This model responds to government concerns over data quality issues, including data provenance, accuracy, and completeness. The data curation model is described in context of a university campus case study, and lessons learned are synthesized for future application in a municipal government context. This research can clarify important issues on trust, quality and accuracy of contributed content, and the role of technology in enhancing or restricting the participation of citizens in the fundamental task of data creation and editing, and how those contributions can enable improved government decision-making.

**Acknowledgments:** This research is supported through contributions from Esri Canada and applies existing Esri Base Map Feedback Application technology. The author acknowledges the support of Dr. Brent Hall, Esri Canada, in developing this conceptual framework.

## References

- Brabham, D. C. (2009). Crowdsourcing the Public Participation Process for Planning Projects. *Planning Theory*, 8(3), 242–262. doi:10.1177/1473095209104824
- Dodge, M., & Kitchin, R. (2013). Crowdsourced cartography: mapping experience and knowledge. *Environment and Planning A* doi:10.1068/a44484
- e Silva, A., & Frith, J. (2010). Locative mobile social networks: Mapping communication and location in urban spaces. *Mobilities*, 5(4), 485–505.
- Elwood, S., Goodchild, M. F., & Sui, D. (2012). Prospects for VGI Research and the Emerging Fourth Paradigm. In *Crowdsourcing geographic knowledge* (pp. 361–375). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-4587-2\_20

- Feick, R., & Roche, S. (2012). Understanding the Value of VGI. In *Crowdsourcing geographic knowledge* (pp. 15–29). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-4587-2\_2
- Ganapati, S., & Reddick, C. G. (2012). Government Information Quarterly. *Government Information Quarterly*, 29(2), 115–122. doi:10.1016/j.giq.2011.09.006
- Girres, J., & Touya, G. (2010). Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, 14(4), 435–459.
- Goetz, M., & Zipf, A. (2012). The Evolution of Geo-Crowdsourcing: Bringing Volunteered Geographic Information to the Third Dimension. In *Crowdsourcing geographic knowledge* (pp. 139–159). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-4587-2\_9
- Gurstein, M. B. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, 16(2).
- Haklay, M., & Ellul, C. (2011). Completeness in volunteered geographical information—the evolution of OpenStreetMap coverage in England (2008–2009). *Journal of Spatial Information Science*, (0), In revision.
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268. doi:10.1080/10580530.2012.716740
- Johnson, P. A., & Sieber, R. E. (2012a). Motivations driving government adoption of the Geoweb. *GeoJournal*, 77(5), 667–680.
- Johnson, P. A., & Sieber, R. E. (2012b). Situating the Adoption of VGI by Government. In D. Sui, S. Elwood, & M. Goodchild (Eds.), *Crowdsourcing geographic knowledge* (pp. 65–81). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-4587-2\_5
- McDermott, P. (2010). Building open government. *Government Information Quarterly*, 27(4), 401–413. doi:10.1016/j.giq.2010.07.002

- Meijer, A. J., Curtin, D., & Hillebrandt, M. (2012). Open government: connecting vision and voice. *International Review of Administrative Sciences*, 78(1), 10–29. doi: 10.1177/0020852311429533
- Onsrud, H., & Pinto, J. (1991). Diffusion of geographic information innovations. *International Journal of Geographical Information Science*, 5(4), 447–467.
- Sui, D., & Delyser, D. (2011). Crossing the qualitative–quantitative chasm I: Hybrid geographies, the spatial turn, and volunteered geographic information (VGI). *Progress in Human Geography*.
- Tinati, R., Carr, L., Halford, S., & Pope, C. (2012). Exploring the Impact of Adopting Open Data in the UK Government. *Digital Futures 2012*, 3.
- Yu, H., & Robinson, D. (2012). The New Ambiguity of “Open Government,” 1–31. Retrieved from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2012489](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2012489)