

Who Are Government OpenData Infomediaries? A Preliminary Scan and Classification of Open Data Users and Products

Peter A. Johnson and Sarah Greene

Abstract: *Open data, that is, the provision of government data in a publicly accessible, machine-readable format, with liberal usage terms, has become commonplace. Despite the promise of open data, there are many questions about who is accessing government open data and what they are using it for. This research presents a characterization of the infomediary, a third party who accesses government open data and creates value-added products from it. Using four major Canadian municipal open data providers, we present an information scan and classification of open data infomediaries and the products they create. Five classifications of infomediary are proposed: government, private sector, NGO, academic, and media. Within each of these classifications, the type of infomediary products created and the delivery method used are summarized. Findings from this research indicate a diversity in infomediary actors and products, but that this activity is largely concentrated in government and private-sector infomediary types. Further considerations of the impact of infomediary activity on government open data provision are presented as important future directions of research.*

INTRODUCTION

In recent years, many governments across North America have begun to provide open data. This release of government information to the public, with minimal use restrictions, is used to enhance transparency and accountability of government, and to drive creative reuse of government data (Bedini et al. 2014; Janssen, Charalabidis, and Zuiderwijk 2012). Despite these accessibility-related motivations for data release, much of the open data currently published by governments is provided in a specialist format, annotated with government-specific terms, tags, or metadata, and often provided in “bulk,” covering a wide time frame or geographic area (Conradie and Choenni 2014, Sieber and Johnson 2015). This method of open data provision reflects the still-emerging mandate of governments as data providers to the public, with limited focus on the technical issues that such users may encounter. Rather, open data provision is based around government exposing data that often is found in the same format and structure as would be used internally. As a result, it can be challenging for nontechnical end-users to access, analyze, and apply open data to projects (Janssen and Zuiderwijk 2014; Zuiderwijk, Janssen, and Dwivedi 2015). This disconnect between the data creation and use by a specific set of users inside of government and its current sharing to a wider, more diverse user base outside of government can potentially limit the overall value of open data. Converting government open data into accessible information that can lead to action can necessitate a third party—the *infomediary*. An infomediary is an entity that takes open government data, in whatever format it may be provided in, and makes it more accessible and useful for end-users, through added analysis, combination with other data sources, and through visualization or publication (Janssen and Zuiderwijk 2014). For

example, journalists act as infomediaries when they access government open data to write a story for publication. Also, community groups, NGOs, and private companies can act in this way, taking raw data, adding value through analysis, and then communicating these results to a specific audience. In this way, the value of open data is not limited to its simple provision, but rather in how use is made of it (Bruin, Bregt, and Ven 2001; Janssen et al. 2012).

A robust knowledge of infomediaries and their role in translating open data to actionable information is only just now emerging from academic literature, including a characterization of the main types of infomediaries, how they transform government data, and the uptake of services that they provide for others further along the chain of data use. The goal of this research is to present an overview of the types of infomediaries that access government open data. These categories of infomediaries include government agencies, the private sector, nongovernmental organizations and community groups, academics and researchers, and the media. This paper draws on an information scan of four leading Canadian municipal-government open data portals (Toronto, Edmonton, Vancouver, Ottawa) to identify the publicly available work of open data infomediaries. A classification of the main types and form of open data infomediaries then is created. We use this classification to advance a discussion on the role that infomediaries play in enabling broader access to and enhancement of government open data. This research provides a foundation for further research on the role and functioning of open data infomediaries, and can provide feedback to government open data providers looking to develop their open data user ecosystems.

OPEN DATA AND OPEN DATA INFOMEDIARIES

The provision of open data by government has increased dramatically in recent years. Open data is government data on infrastructure, spending, services, and procedures, made publicly available for free, online, under a permissive-use license (Bedini et al. 2014, Bonina 2013). Government open data portals are an important first step towards achieving some of the promise of open government by sharing data within and across governments, as well as with a broad community of users, including citizens (Bedini et al. 2014; Charalabidis, Alexopoulos, and Loukis 2016). The traditional open data portal model provides open data, online, in a variety of raw data formats. This “data over the wall” model is not necessarily focused towards broad or easy access by average citizens (Sieber and Johnson 2015). This focus on developers or other technically minded users compared to citizens comes despite long-standing pressure for government to provide information in citizen-centric formats (Luna-Reyes, Bertot, and Mellouli 2014).

Infomediaries can be defined as specific categories of open data users who extract, aggregate, and transform data, altering it into a format that is seen as valuable, beneficial, and, most important, usable to the general public (Bonina 2013). Government infomediaries include open data providers as well as other government departments within a municipality that create value-added products with open data, making it more accessible to the end-user (Janssen and Zuiderwijk 2014; Magalhaes, Roseira, and Strover 2013). Infomediaries can play a variety of roles, ranging from increasing transparency and accountability to encouraging innovation and economic growth within a city (Davies and Frank 2013). Generally, an important role of a government infomediary is the ability to visualize complex relationships and information, creating value-added services for other departments within the government, and additionally for citizens (Deloitte Analytics 2012). Using tools for visualization and data analysis, such as the creation of charts, graphs, and applications, improves accessibility to open data for any audience, particularly those without a technical background (Fumega 2014, Graves and Hendler 2013). Technically skilled infomediaries are critical to the success of open data programs, for they have the ability to engage citizens who otherwise would not access open data because of a lack of interest, skills, or resources (Fumega 2014). Infomediaries act to help bridge the gap between the data that is available from government and the potential for it to be used by citizens. This provides an opportunity for a range of end-users to both access and effectively use this data (Chattapadhyay 2014, Davies 2014, Fumega 2014).

G4 CITIES AND METHOD

Open data provision in Canada has expanded dramatically over the past decade. Starting as an emerging phenomena at only a select few cities, open data now is commonly provided by many municipal and provincial governments. Many of those municipal governments offering open data do so using the Canadian Open Government License, developed by the federal Canadian govern-

ment, which provides for attribution-only, mixed commercial/noncommercial use of open data (Roy 2014). In parallel to this increased provision of open data by governments, private-sector use of open data in Canada is growing. The recent launching of the Open Data 150 found more than 150 private-sector companies using Canadian open data for a variety of purposes (GovLab 2017). Given this increase in both provision and use of open data, we investigate the current range of infomediary activities in major Canadian open data-providing cities, using a targeted review of publicly available information.

Taking the “G4,” the most developed open data-providing Canadian cities of Edmonton, Toronto, Ottawa, and Vancouver as our study subjects, we conducted a Web scan of municipal Web sites, including searching their open data catalogs, available galleries of finished products, results from hackathons or other coding contests, and other departmental pages to gather examples of infomediaries. Data for this research was collected from September to December of 2015. Additionally, targeted key searches were undertaken to find examples of products created by a specific infomediary category in a G4 city. These targeted searches included some key terms such as *Toronto AND open data* along with more specific searches such as *Vancouver AND data journalists*. Results then were compiled to reveal both general results of the types of infomediary products created and the open data used, and, additionally, a profile of each infomediary type was created, showing commonalities and differences among the G4 cities. Though an exhaustive record of all open data use by infomediaries is not possible, this research gives a starting point for the classification of categories of infomediary, an overview of the types of products created from major municipal open data catalogs, and how these products are delivered to end-users.

CLASSIFICATION OF INFOMEDIARIES, INFOMEDIARY PRODUCTS, AND DELIVERY METHOD

Infomediary products for the G4 cities were classified in three ways: infomediary type, infomediary product, and delivery method. First, the classification of infomediary type drew from existing literature on open data-user communities, most notably Bonina’s (2013) overview of open data business models, Magalhaes et al.’s (2013) framework of open-government intermediaries, and Janssen and Zuiderwijk’s (2014) classification of infomediary business models. This literature, plus other works, describes key open data-user communities as including government itself (Kankanhalli, Zuiderwijk, and Tayi 2017), the private sector (Bonina 2013), journalists and media organizations, researchers and academics, and nonprofits or community groups (Chan, Johnson, and Shookner 2016; Safarov, Meijer, and Grimmelhuisen 2017). Given the identification of these user communities, we adopt a similar classification to frame this research, though we also acknowledge that a complete census of all open data users is difficult to track (Johnson 2016).

Table 1. Breakdown of infomediary type by city

	Edmonton	Ottawa	Toronto	Vancouver	Total	Percent
Category						
Government	29	13	12	8	62	27
Private Sector	18	49	33	26	126	55
NGO	1	3	6	4	14	6
Academic	2	3	1	1	7	3
Media	5	0	11	3	19	8
Total	55	68	63	42	228	100
Percent	24	30	28	18	100	

Table 2. Type of product by infomediary type

	Gov	%	Private Sector	%	NGO	%	Media	%	Academic	%	Total	%
Searchable Database	20	32	64	51	7	50	0	0	0	0	91	40
Interactive Map	23	37	40	32	1	7	0	0	2	29	66	29
Reporting Service	2	3	1	1	0	0	0	0	0	0	3	1
Find a Service	2	3	15	12	0	0	0	0	0	0	17	7
Report	5	8	0	0	1	7	0	0	1	14	7	3
Static Map	6	10	2	2	1	7	0	0	0	0	9	4
Graph	4	6	2	2	2	14	0	0	1	14	9	4
Game	0	0	1	1	0	0	0	0	0	0	1	0
Web Page	0	0	1	1	0	0	0	0	0	0	1	0
Articles	0	0	0	0	2	14	19	100	3	43	24	11
Total	62	100	126	100	14	100	19	100	7	100	228	100
%	27		55		6		8		3		100	

To frame this research, we used five main types of infomediaries: government, private sector, NGO and/or community groups, academic researchers, and journalists/media. Table 1 shows a breakdown of infomediary type by each of the cities included in this research. These broad categories serve as a high-level typology of the open data infomediary role. Infomediary product categories were created through aggregating similar-type products found at the surveyed cities. Table 2 presents these types of products by infomediary type. Though we classified ten different types of infomediary products, significant numbers are found only in four types: searchable database (such as a transit, recreation, or festival schedule), interactive map (using maps to display a user-selected range of information, such as municipal infrastructure), articles that contain data visualizations (prepared for mass media, blogs, etc., containing charts, tables, maps), and find-a-service portals, such as locating street food vendors or historic buildings. Lastly, delivery method of infomediary products was classified into one

of four categories: database (a user-query dataset, but with limited visualization or analysis capabilities), a Web site (publicly posted document that may contain text, images, maps, and information derived from a database), mobile app (similar to a Web site but designed specifically for a telephone interface), and a specific computer application to run on a desktop or laptop computer. Delivery method was used to track the split between products provided in specialist formats, broadly on the Web, and as a mobile app, targeting mobile telephone users. Table 3 gives a breakdown of each of these delivery methods based on infomediary type, as well as links to examples. These results are presented within each infomediary type, across all surveyed cities, with breakdowns of type of product and delivery method. These latter two areas (products and delivery method) are described in context of each of the high-level infomediary types—government, private sector, NGO, academic, and media.

GOVERNMENT INFOMEDIARIES

Government infomediaries include open data providers themselves, governments at the same or alternate level (municipal, state/province, federal), as well as departments within a given government that create different value-added products using government open data (Janssen and Zuiderwijk 2014; Jetzek, Avital, and Bjørn-Andersen 2013). From our study sample, 27 percent of all the examples of infomediary activity gathered across the G4 cities were produced by government infomediaries (see Table 1). This was the second largest category of infomediary, after the private sector, though we discuss this infomediary first, for it is the closest to the originating source of data. With the delivery of open data providing a foundation for further actions, other government departments or agencies act as infomediaries through the analysis of data and resulting production of more accessible content directed towards a specific range of end-users (Deloitte Analytics 2012). This value-added creation can take the form of visualization or data analysis, such as the creation of interactive maps, charts, graphs, and applications that improve the accessibility to open data for many audiences, particularly those without a technical background (Fumega 2014, Graves and Hendler 2013).

Table 2 provides a breakdown of the main types of products made by government infomediaries. For example, across all G4 cities, the use of interactive maps (37 percent) was a common way to allow users to display, filter, and perform basic analysis. Searchable databases (32 percent), such as a listing of recreation programs or transit times, also were commonly seen types of products. Additionally, the creation of static maps (10 percent) highlighting such city services as bicycle routes was seen, as well, as a less common but still consistent form of sharing open data information. Creating visualizations for reports (8 percent) through specialized data-management software, such as Tableau (<http://www.tableau.com/>) also were seen throughout the G4 cities, particularly for sharing complex budget information through reports that relied heavily on simple graphics including charts and graphs. An example of this is the city of Edmonton's Citizen Dashboard (<https://dashboard.edmonton.ca/>), an online tool designed to engage nontechnical end-users through easier-to-understand visualizations of municipal performance indicators. These infomediary products were delivered largely as Web sites (75 percent) and as mobile apps (18 percent), showing government infomediaries as focusing on wider distribution via the Web, compared to more fragmented mobile operating systems (see Table 3). Overall, across the G4 cities, government infomediaries were critical players in working with government open data to create value-added products.

PRIVATE SECTOR

The private sector is considered to drive demand for open data through the potential for economic growth and overall ability to innovate and create services for citizens (Bonina 2013, Davies 2014, Gray and Darbishire 2011). Private-sector infomediaries

include companies and individual developers who use open data to create new products (Deloitte Analytics 2012). In particular, the creation of mobile applications has been seen as a major outlet in which the products of open data can be shared widely (Yang and Kankanhalli 2013). Across the study sample, private-sector companies represented 55 percent of all infomediaries found in the G4 cities (shown in Table 1). This is the largest category of infomediary, with more than twice as many examples as the next largest category (government).

Previous research has found that open data is used by private-sector developers to design Web sites and maps and build Web-based or mobile applications, which may include real-time information (Davies and Edwards 2012, Deloitte Analytics 2012) into research, to operational project statistics. A common example of a product created by private-sector developers is a public-transit application. These applications use real-time bus Global Positioning System (GPS) information to help citizens interact more efficiently with public transit (Rojas 2012). The sample of open data products created by the private sector using G4 city data is dominated by three main types of products: searchable databases (51 percent), interactive maps (32 percent), and find-a-service (12 percent) (see Table 2). Searchable databases created by private-sector infomediaries included transit routing apps. With bus schedule data available to the public in all G4 cities, many developers have used this data to create transit applications that show bus routes and allow for trip planning. Interactive maps created by private-sector infomediaries include real-estate applications that allow users to filter demographic and neighborhood amenity information along with for-sale listings. Find-a-service applications include third-party apps that facilitate interaction with data such as the location of public-access defibrillators (Ottawa) or festival events (Toronto). Overall, the types of products created by private-sector infomediaries is very diverse. Another notable product created by a private-sector infomediary includes the samples-only game created using open data of the cities of Ottawa, Toronto, and Vancouver. This Web-based interactive game titled "Click that 'hood'" (<http://click-that-hood.com/>) engages citizens in learning the names of all the neighborhoods located within a city. Although this was the only game found in the research, it shows the potential of open data to be used outside of the realm of just informing or guiding citizens through a city, but additionally engaging citizens in enjoyable activities.

Private-sector infomediaries primarily delivered their open data-based products as mobile applications (62 percent, shown in Table 3). Many of these applications were creations from hackathon events, hosted by the G4 cities. The private-sector companies creating mobile applications therefore ranged from one-off applications by individual developers to a number of applications made by larger public companies. This significant use of the mobile application for delivering open data products shows the integration of private-sector infomediaries with the more easily monetized mobile-app ecosystem, compared to a Web site or other method.

NONGOVERNMENTAL ORGANIZATIONS (NGO)/ COMMUNITY GROUPS

Nongovernmental organizations (NGOs) and community groups are infomediaries who produce a wide range of products with open-government data (Chattapadhyay 2014). An important role played by this type of infomediary includes encouraging government accountability and effectiveness (Davies 2014). NGOs provide key insights into which datasets should be opened to help produce high value and engagement with the public (Ubaldi 2013). Another task performed by NGOs includes the creation of reports, most of which focus on highlighting education, health, and other societal issues (Graves and Hendler 2013). NGOs additionally focus on sharing budget-related information in easy-to-understand formats (Mejabi, Azeez, Adedoyin, and Oloyede 2014). An important function of NGO and community-group infomediaries are their ability to collaborate with governments, building trust in opening information to the public, and additionally opening data to a wider audience of potential users, including grassroots communities (Gray and Darbishire 2011). Although NGOs are seen as a smaller group of infomediaries for open data, they are growing in number and importance through their engagement with open data in relation to domestic priorities and issues (Davies 2014).

NGO and community-group infomediaries represented 6 percent of the total number of all infomediary products made with G4 city data (shown in Table 1). Additionally, for NGO and community-group infomediaries, the sample size is very small, with only 14 individual products recorded across the four cities surveyed. Infomediary products produced by NGOs in the G4 cities were dominated by searchable databases (50 percent), with only a handful of singular instances of other categories. Delivery methods used to provide these products were dominated by Web sites (57 percent), a sharp contrast to the mobile app-heavy provision of the private-sector infomediaries (see Table 3). This contrasts to literature on the role of NGOs and nonprofits in accessing and using open data to create products of wider value to citizens. Despite low individual numbers of infomediary products, there were notable examples that could have broad use by citizens and others. For example, in the city of Vancouver,

a collaborative project was created called Urban Opus (<http://urbanopus.net/>), an organization that brings together community organizations, governments, and private-sector corporations that helps to develop original data-driven services to engage civic audiences. Similarly, Geographic and Effective Measures Services (<http://www.gems-spc.ca/>) is another consulting NGO located in Ottawa that creates profiles of given neighborhoods, as well as custom printing maps. Lastly, the rePresent API, a project of Open North (<https://represent.opennorth.ca/>), uses a range of open data to provide a platform for informing citizens about their local political representatives.

Academics/Research

Research infomediaries include academics and consultants that use open data to create academic papers or reports, analyzing topics related to government open data (Davies 2010). This group of infomediaries analyzes open data and provides the results of the analysis through both text and visualizations. The purpose of research infomediaries is to use open data to solve important issues and provide insights into government actions and services. Commonly, transit ridership was analyzed to determine the success of transit systems in different municipalities. Open data additionally can be used alongside other datasets to enhance research and potential results. Research infomediaries are of particular importance for their products include value added to the original data, providing further information and analysis of governments to citizens. Research infomediaries in many cases perform secondary research, using government open data to cross-check findings, provide new findings, and generally inform citizens about government-related information (Davies and Edwards 2012). Across the G4 cities sample, research infomediaries were found to represent only a small number of the totally number of infomediary examples (3 percent, see Table 1). Given this small total number of examples, it is difficult to draw any meaningful conclusions about the specific types of infomediary products created, though of those examples, 57 percent were articles and 29 percent were interactive maps (shown in Table 2). It is notable that all the research infomediary products were delivered using the Web page method (see Table 3).

Table 3. Product delivery method by infomediary type

	Gov	%	Private Sector	%	NGO	%	Media	%	Academic	%	Total
Database	5	8	0	0	3	21	0	0	0	0	8
Web Site	46	74	42	33	8	57	19	4	7	0.5	122
Mobile App	11	18	78	62	3	21	0	0	0	0	92
Computer App	0	0	6	5	0	0	0	0	0	0	6
	62	100	126	100	14	100	19	4	7	0.5	228

Media

Primarily titled data journalists, media infomediaries access open data, analyze the information, and provide citizens with results and analysis of this information. With the increase in social media, more and more venues are opening up in which data journalists can share reports on open data, allowing citizens to discuss and comment on the information provided. The introduction of wider structures of media networks, such as social media, has positively impacted the availability and sharing of open data to citizens (Davies 2010). With the rise of open data, there has been a creation of a new group of journalists known as data journalists. Data journalists act as infomediaries by accessing government open datasets to identify stories and share information with the general public (Mejabi et al. 2014, Roy 2014, Sapkota 2014). The role of data journalists is seen as fundamental in distributing government data and engaging citizens in government open data (Fumega 2014). Particularly in developing countries and/or communities, data journalists are seen as important infomediaries for they are most accessible to citizens, including those who may not have access to or understand technologies such as the Internet (Mejabi et al. 2014, Sapkota 2014). Across the G4 cities sample, media represented a small total number of instances of infomediary products, with 8 percent of the total amount (see Table 1). Despite this small number of specific instances, there is potential for a large audience or readership with media infomediary products. Overall, media infomediaries created articles, with 100 percent of the infomediary products being articles (see Table 2). These articles were uniformly distributed via Web sites (shown in Table 3). Blog posts were a common form of journalism including government open data. Many of these blog posts include tutorials on how to use and access open data, teaching citizens about the data they are using.

A CLASSIFICATION OF OPENDATA INFOMEDIARIES AND “WEBS” OF OPEN DATA VALUE

Through this characterization of the Canadian G4 cities open data infomediaries, several key themes have emerged. First is that there is a diverse range of organizations, whether government agencies, private-sector companies, media, researchers, or non-profits and community groups that access government open data. These broad categories of infomediary each access data for a variety of purposes and use open data in various ways related to their organizational and individual goals. Despite this broad range of actors that access open data, the infomediary landscape is dominated by private-sector and government infomediaries. As characterized, the types of products created through the use of open data by infomediaries are varied and can include databases, Web sites, mobile apps, blog posts, reports, and more. Table 2 provides an overview of these uses of open data and also the ways in which they are distributed, either through Web sites, mobile apps, or dedicated conventional computer programs. Given this

range and diversity of intervening actors and end-products created, a significant outcome of this work is to suggest a refocusing of current attention on open data from the simple provision of open data to the questions of how other entities are accessing this data to create downstream value. The identification of specific users, each with unique use cases, can be used to better understand the impact of data provision by government. Within open data literature and practice, there is a significant focus on enabling access to raw data from government (Denis and Goëta 2014, Sieber and Johnson 2015). This literature focuses largely on topics of data format, licensing and copyright, as well as organizational issues that may impede government provision of open data (Chattopadhyay 2014, Harvey 2007). These concerns are typical of early innovation life-cycle issues (Rogers 2010) where technical or mechanical issues, those of just providing data (Johnson and Sieber 2013, Sieber and Johnson 2015), are paramount. As open data matures and becomes more standardized across governments, this discussion moves naturally on to one of value and utility of the data access provided. It is to this growing field of research that this work aims to contribute, through the classification of the infomediary class of open data users.

Building on the classification of open data infomediaries presented here, we propose that infomediaries are a primary way that open data creates a “web” of value, with a single point of access by an infomediary creating broader access through value-added developments driven by third parties. This concept is supported through recent studies from Janssen (2012) and the New York University GovLab (Verhulst and Young 2016), both of which aim to evaluate the impact of open data. Janssen (2012) indicates that open data itself has little value, and that this value is only realized after use. Publication of data in an open format for easy access is simply the first step towards use and impact; however, for this impact to be generated, infomediary action is required. Impact of open data is considered by GovLab (Verhulst and Young 2016) as the product of a series of enabling conditions, including the leveraging of open data by partnerships both within and external to the data-generating government. This presentation of partnerships as a key to unlocking the value of open data is a finding that is revealed in this work as well, with the infomediary class of open data users representing a manifestation of the partnership or open contract (whether explicit or implicit) between the data-generating government and the user (Verhulst and Young 2016). This contract includes not only the ability to access data, but to enhance, manipulate, and, in turn, share open data to other downstream users, forming the “web” of value where impact, generated through initial sharing between government and infomediary, is enhanced through infomediary activity. For example, in the survey of G4 cities, infomediaries often combined datasets (both open and closed) and enhanced open datasets to create a value-added proposition for downstream users. Through the sharing of data in this fashion, impact also is diversified, moving into areas that may not have been initially considered by the data owners. An example of this is the aforementioned “Click that ‘hood’” game, where the creative use of neighborhood boundary

data has potential to increase civic pride and knowledge of an area, an outcome that is vastly different than the likely intended goal of providing neighborhood boundary data to facilitate record keeping, map making, and other typical uses.

Despite the potential formation of webs of open data value by infomediary actors, many questions remain. First, as revealed by this research, products created by government and private-sector infomediaries vastly outnumber those created by other infomediary groups. This may indicate that these two categories of infomediary are the most significant users of government open data; however, this may ignore the extent to which these products actually are used. For example, though the private sector may create a number of products from government open data, how many of these products are essentially duplicates of one another, with minor differences? And how many of these products generate a robust number of users? These relevant questions indicate that in many ways, tracking the value of open data requires the tracing of all nodes on the data access and use web, following through from transformations, analyses, or unique data combinations that may be introduced by infomediaries. Specific numbers or quantitative measures of use, such as downloads of datasets, number of products created, or even the number of viewers of a particular news story partially made with open data, while straightforward to quantify, lack in understanding of the impact and role played by open data, particularly in further downstream uses of products created by primary infomediary activity. Though a finding of this research is that the open data infomediary space is dominated by government and private-sector actors, this should not be extended to include that these infomediaries necessarily lead to a greater impact of open data via these activities.

CONCLUSIONS

This research provides a characterization of open data infomediaries, using four major Canadian cities as sources for comparing the types of infomediaries, products produced, and delivery methods used. Infomediaries were framed into five distinct categories, based on previous research and sectoral characteristics: government, the private sector, NGO or community organizations, academics or researchers, and the media. The total number of instances of infomediary products across the four cities was dominated by the private-sector and government categories of infomediary. Products produced by infomediaries depended largely on the type of infomediary, but typically focused on searchable databases, interactive maps, find-a-service tools, and media/blog articles. Delivery methods were dominated by Web sites and mobile applications. A significant message from this work for governments that provide open data is that enabling access to data, though clearly important, is only the first step in generating impact from open data. Providing access to data is simply the beginning of what can be considered a “web” of open data use that is driven by infomediary actors. As noted by Janssen et al. (2012), the simple provision of open data has little value. It is from the process of use, via numerous infomediaries, that value is created.

It is important to consider that all open data programs are

not identical, and that the individual model of open data provision may support or restrict infomediary activity, or even favor certain types of infomediaries. For example, Sieber and Johnson’s (2015) models of open data provision indicate some forms of data provision would better support private-sector infomediary activity, most notably the “data over the wall” model. Other models, such as “code exchange” and “civic issue tracker,” more deeply involve government in creating data endpoints for citizen use of open data (Sieber and Johnson 2015). Lastly, the prospect of participatory open data is a model where citizens and government co-create products with open data. This model would firmly cement the role of government as an infomediary when dealing with citizens, with deep responsibility not only for the data itself, but in how it gets used.

Many challenges exist to the use of open data, and infomediaries accessing and using data many encounter many technical, organizational, and social challenges. Technical challenges include differences in data collection, format, licensing, and varying levels of data completeness and quality that may impede use (Johnson, Sieber, Scassa, Stephens, and Robinson 2017). For government to support the use of open data by infomediaries, there are many changes to data provision that could ameliorate or reduce these challenges. First, the development and enforcement of open data standards (Goëta 2014, Piovesan 2015) can provide infomediaries with more seamless data that crosses jurisdictions. This can support the use and integration of multiple datasets from across different governments. An agreed-upon set of standards for the provision of open data also can ensure that infomediaries can access data in a uniform fashion, reducing development costs generated through working with and processing data in a variety of formats.

At a higher level, the provision of open data by governments to a range of infomediaries, some of which may be building billion-dollar business models using open data, raises a number of questions. First, what is the role of government in supporting, for free, these types of corporate endeavors (Johnson et al. 2017)? From the business-development side, can businesses rely on government to continue providing, collecting, and licensing data for free, or could cost-recovery fees be reintroduced to support government data-collection and management efforts (Sieber and Johnson 2015)? Lastly, what priority will government place on the development and provision of data that meets a corporate need versus data that meets a societal need? Could a priority be placed on that data that is mission-critical for private-sector infomediaries (with teams of lawyers and effective lobbyists), compared to data that may have little commercial value, but instead be of broader social value, or political value, such as transparency or procedural data? These are critical next-step questions that will further refine the role of the infomediary and its evolving relationship with the data-generating government.

Ultimately, the development of a class of open data infomediaries indicates that data-provisioning governments have many different user communities to work with, and that there may be challenges in meeting the potentially divergent needs of

these communities. A particularly notable challenge is balancing the technical needs of developers versus the nontechnical needs of more general users. With the development of infomediaries, government no longer can simply provide data, but rather must be responsive to the needs of infomediaries who may request standardization of formats, specific datasets required to meet their needs, and further support from the data custodians. This issue, of access to technology by those with varying levels of skill to use a specific piece of software, has been noted in early work on diffusion of GIS (Onsrud and Pinto 1991) and the development of public participation GIS (PPGIS) approaches (Sieber 2006). In contrast with these works, which largely looked at access on an individual or group level, the role of an infomediary organization can be to overcome some of these barriers to access of technology and data that are encountered by society at large. Though the specific functioning of the infomediary as an open data “chauffeur,” helping to spread access and skills in open data analysis and application, is beyond the scope of this work, it remains an important line of questioning, and one that can draw on GIS implementation literature more strongly (Ghose 2001, Sieber 2001).

This research presents an initial view of the role of infomediaries in generating value from government open data within a convenience sample of cities—the G4 Canadian open data cities of Toronto, Vancouver, Edmonton, and Ottawa. As this represents a small subset of all open data—providing cities, this work should be considered exploratory. Additionally, as the total number of infomediaries using government open data cannot entirely be known, there is the possibility that the relative breakdowns of open data infomediary type and product presented here are a preliminary view. Given the emerging state of research on the nature and output of open data users, there still are important lessons to guide future research. Most notably, that government is a major user of open data shows that the infomediary role is broad, and that government has a role to play in open data use, not just provision.

The benefits of open data may take time to emerge, as private-sector users usually look for market opportunities rather than observing the open data to invent new products immediately (World Wide Web Foundation 2015). Although the impacts and benefits of open data may not be present at first, defining them is important in discovering the importance of open data programs (Höchtel, Davies, Janssen, and Schieferdecker 2014). Understanding the relationship between those who publish the data, those who create products based on the open information, and the final end-user is where many of the overall impacts can be found. As many citizens themselves may not be able to interact directly with open data, the role of infomediaries becomes important in determining the overall success of an open data program. Now that many of the initial challenges in providing open data have been surmounted, the development of measures of open data value are taking priority in the research and practitioner fields. As open data moves from a disruptive phenomenon to a commonly found, and in many ways, expected service, there is

increased pressure to justify the continued expense and effort of data provision. In this way, the development of metrics of value are a key component to creating a picture of open data use that includes diverse end-users and also those infomediaries who reach a wider diversity and extent of end-users.

Acknowledgments

This work was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC) Partnership Grant Geothink.

About the Authors

Dr. Peter A. Johnson is an associate professor in the Department of Geography and Environmental Management at the University of Waterloo. His research expertise is in the application and evaluation of geospatial technologies, including open data, the geospatial Web, mobile devices, crowdsourcing, participatory geomatics, and volunteered geographic information.

Corresponding Address:
Department of Geography and Environmental Management
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada
peter.johnson@uwaterloo.ca

Sarah Greene completed her undergraduate degree in geography and environmental management at the University of Waterloo in 2016. She currently is attending Ryerson University, completing a Master’s of Spatial Analysis. Her research interests include studying the impacts and use of open data.

Corresponding Address:
Department of Geography and Environmental Management
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
sagreene@uwaterloo.ca

References

- Bedini, I., F. Farazi, J. Pane, I. Tankoyeu, D. Leoni, and S. Leucci. 2014. Open government data: Fostering innovation. *JeDEM: eJournal of eDemocracy and Open Government* 6(1). Retrieved from <https://www.w3.org/2013/share-psi/wiki/images/e/eb/Feroz.pdf>.
- Bonina, C. M. 2013. New business models and the value of open data : Definitions, challenges and opportunities (August). Retrieved from <http://www.nemode.ac.uk/wp-content/uploads/2013/11/Bonina-Opendata-Report-FINAL.pdf>.
- de Bruin, S., A. Bregt, and M. van de Ven. 2001. Assessing fitness for use: The expected value of spatial datasets. *International Journal of Geographical Information Science* 15(5): 457-71. [Http://doi.org/10.1080/13658810110053116](http://doi.org/10.1080/13658810110053116).
- Chan, M., P. Johnson, and M. Shookner. 2016. Assessing the use of government open data and the role of data infomediaries. The case of Nova Scotia's community counts program. *JeDEM: eJournal of eDemocracy and Open Government* 8(1): 1-27.
- Charalabidis, Y., C. Alexopoulos, and E. Loukis. 2016. A taxonomy of open government data research areas and topics. *Journal of Organizational Computing and Electronic Commerce*. [Http://doi.org/10.1080/10919392.2015.1124720](http://doi.org/10.1080/10919392.2015.1124720).
- Chattapadhyay, S. 2014. Opening government data through mediation: Exploring the roles, practices and strategies of data intermediary organizations in India. *Practices and Strategies of Data Intermediary Organizations in India* (December 31, 2014). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2549734.
- Conradie, P., and S. Choenni. 2014. On the barriers for local government releasing open data. *Government Information Quarterly* 31: S10-17. [Http://doi.org/10.1016/j.giq.2014.01.003](http://doi.org/10.1016/j.giq.2014.01.003).
- Davies, T. 2010. Open data, democracy and public sector reform: A look at open government data use from data.gov.uk (Master's dissertation). University of Oxford. Retrieved from <http://www.opendataimpacts.net/report/>.
- Davies, T. 2014. Open data policies and practice: An international comparison (SSRN Scholarly Paper No. ID 2492520). Rochester, NY: Social Science Research Network, 1-26. Retrieved from <http://papers.ssrn.com/abstract=2492520>.
- Davies, T., and D. Edwards. 2012. Emerging implications of open and linked data for knowledge sharing in development. *IDS Bulletin* 43(5): 117-27.
- Davies, T., and M. Frank. 2013. "There's no such thing as raw data." Exploring the sociotechnical life of a government dataset. In *Web science*. Paris: 2013. Retrieved from http://www.opendataimpacts.net/wp-content/uploads/2013/05/paper_11.pdf.
- Deloitte Analytics. 2012. Open growth—stimulating demand for open data in the UK. Retrieved August 5, 2014, from http://www.deloitte.com/view/en_GB/uk/market-insights/deloitte-analytics/open_data/bfb570a79416b310VgnVCM1000003256f70aRCRD.htm.
- Denis, J. and S. Goëta. 2014. Exploration, extraction and "ra-wification." The shaping of transparency in the back rooms of open data (SSRN Scholarly Paper No. ID 2403069). Rochester, NY: Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=2403069>.
- Fumega, S. 2014. City of Buenos Aires open government data initiative, 1-43. Retrieved from <http://opendataresearch.org/sites/default/files/publications/Final-Opening%20Cities%20-%20Buenos%20Aires%20final%20report.pdf>.
- Ghose, R. 2001. Use of information technology for community empowerment: Transforming geographic information systems into community information systems. *Transactions in GIS* 5(2): 141-63. [Https://doi.org/10.1111/1467-9671.00073](https://doi.org/10.1111/1467-9671.00073).
- Goëta, S. 2014. The daily shaping of state transparency: Emerging standards in open government data. In *4S Conference*. Buenos Aires, 1-12.
- GovLab. 2017. Open data 150 Canada. Retrieved from <http://canada.opendata500.com>.
- Graves, A., and J. Hendler. 2013. Visualization tools for open government data. In *Proceedings of the 14th Annual International Conference on Digital Government Research*. New York: ACM, 136-45. [Http://doi.org/10.1145/2479724.2479746](http://doi.org/10.1145/2479724.2479746).
- Gray, J., and H. Darbishire. 2011. Beyond access: Open government data and the right to (re)use public information (January). Madrid: Access Info.
- Harvey, F. 2007. Just another private-public partnership? Possible constraints on scientific information in virtual map browsers. *Environment and Planning B: Planning and Design* 34(5): 761-64.
- Höchtel, J., T. Davies, M. Janssen, and I. Schieferdecker. 2014. Open data: Growing up and getting specific. *JeDEM: eJournal of eDemocracy and Open Government* 6(1): 2014. Retrieved from <http://repository.tudelft.nl/view/ir/uuid:e9a68745-e209-41b6-9037-0ebe438d642d/>.
- Janssen, M., Y. Charalabidis, and A. Zuiderwijk. 2012. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management* 29(4): 258-68. [Http://doi.org/10.1080/10580530.2012.716740](http://doi.org/10.1080/10580530.2012.716740).
- Janssen, M., and A. Zuiderwijk. 2014. Infomediary business models for connecting open data providers and users. *Social Science Computer Review* 32(5): 694-711. [Http://doi.org/10.1177/0894439314525902](http://doi.org/10.1177/0894439314525902).

- Jetzek, T., M. Avital, and N. Bjørn-Andersen. 2013. Generating value from open government data. In the 34th International Conference on Information Systems. ICIS 2013. Retrieved from <http://forskningbasen.deff.dk/Share.external?sp=S83326cef-b316-4aa7-87d9-6049ff3235b7&sp=Scbs>.
- Johnson, P. A., and R. E. Sieber. 2013. Situating the adoption of VGI by government. In D. Sui, S. Elwood, and M. Goodchild, Eds. *Crowdsourcing geographic knowledge*. Dordrecht: Springer Netherlands, 65-81.
- Johnson, P. A. 2016. Reflecting on the success of open data: How municipal government evaluates their open data programs. *International Journal of E-Planning Research* 5(3): 1-12.
- Johnson, P. A., R. Sieber, T. Scassa, M. Stephens, and P. Robinson. 2017. The cost(s) of geospatial open data. *Transactions in GIS* 21(3): 434-45.
- Kankanhalli, A., A. Zuiderwijk, and G. K. Tayi. 2017. Open innovation in the public sector: A research agenda. *Government Information Quarterly* 34(1): 84-89.
- Luna-Reyes, L. F., J. C. Bertot, and S. Mellouli. 2014. Open government, open data and digital government. *Government Information Quarterly* 31(1): 4-5. <http://doi.org/10.1016/j.giq.2013.09.001>.
- Magalhaes, G., C. Roseira, and S. Strover. 2013. Open government data intermediaries: A terminology framework. In the Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance. New York: ACM. 330-33. <http://doi.org/10.1145/2591888.2591947>.
- Mejabi, O., A. Azeez, A. Adedoyin, and M. Oloyede. 2014. Case study report on investigation of the use of the online national budget of Nigeria. University of Ilorin. Retrieved from <http://opendataresearch.org/sites/default/files/publications/Investigation%20of%20Open%20Budget%20Data%20in%20Nigeria-print.pdf>.
- Onsrud, H. J., and J. K. Pinto. 1991. Diffusion of geographic information innovations. *International Journal of Geographical Information Systems* 5(4): 447-67. <https://doi.org/10.1080/02693799108927868>.
- Piovesan, F. 2015. Beyond standards and regulations: Obstacles to local open government data initiatives in Italy and France. Université Paris XIII, Università di Torino. Retrieved from https://www.researchgate.net/publication/282861053_Beyond_Standards_and_Regulations_Obstacles_to_Local_Open_Government_Data_Initiatives_in_Italy_and_France_Economic_Policies_in_the_Age_of_Globalization_Knowledge_and_Innovation_Policies.
- Rogers, E. M. 2010. *Diffusion of innovations*. Simon and Schuster.
- Rojas, F. M. 2012. *Transit transparency: Effective disclosure through open data*. Cambridge, MA: New England University Transportation Center. Retrieved from http://www.transparencypolicy.net/assets/FINAL_UTC_TransitTransparency_8282012.pdf.
- Roy, J. 2014. Open data and open governance in Canada: A critical examination of new opportunities and old tensions. *Future Internet* 6(3): 414-32. <http://doi.org/10.3390/fi6030414>.
- Safarov, I., A. Meijer, and S. Grimmelikhuijsen. 2017. Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity* 1-24.
- Sapkota, K. 2014. Research poster: Towards an action research agenda for inclusive open data in Nepal. <http://www.opendataresearch.org/project/2013/freedomforum>.
- Sieber, R. E. 2001. Conforming (to) the opposition: The social construction of geographical information systems in social movements. *International Journal of Geographical Information Science* 14(8): 775-93. <https://doi.org/10.1080/136588100750022787>.
- Sieber, R. 2006. Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers* 96(3): 491-507.
- Sieber, R. E., and P. A. Johnson. 2015. Civic open data at a crossroads: Dominant models and current challenges. *Government Information Quarterly* 32(3): 308-15. <http://doi.org/10.1016/j.giq.2015.05.003>.
- Ubaldi, B. 2013. Open government data: Towards an empirical analysis of open government data initiative (Working Papers on Public Governance No. 22). Paris: OECD. Retrieved from <http://dx.doi.org/10.1787/5k46bj4f03s7-en>.
- Verhulst, S., and A. Young. 2016. Open data impact: When demand and supply meet. GovLab. Retrieved from www.odimpact.org.
- World Wide Web Foundation. 2015. *Open data barometer*, Second Ed. Retrieved from <http://www.opendatabarometer.org>.
- Yang, Z., and A. Kankanhalli. 2013. Innovation in government services: The case of open data. In Y. K. Dwivedi, H. Z. Henriksen, D. Wastell, and R. De', Eds. *Grand successes and failures in IT. Public and private sectors*. Springer Berlin Heidelberg, 644-51. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-38862-0_47.
- Zuiderwijk, A., M. Janssen, and Y. K. Dwivedi. 2015. Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly* 32(4): 429-40. <http://doi.org/10.1016/j.giq.2015.09.005>.