

The Making of a ‘Top’ Open Data City: A Case Study of Edmonton’s Open Data Initiative

Jacqueline Corbett
Université Laval
jacqueline.corbett@fsa.ulaval.ca

Mathieu Templier
Université Laval
mathieu.templier@fsa.ulaval.ca

Hirotooshi Takeda
Université Laval
hirotoshi.takeda@fsa.ulaval.ca

Abstract

In recent years, various models and indexes have been proposed to evaluate and rate the performance of open data initiatives. However, little research examines cities’ open data initiatives in relation to these indexes and how cities achieve open data success. Through an exploratory case study of Edmonton, Canada’s top ranked open data city, this research sheds light on the mechanisms contributing to top-rated and successful open data initiatives. Our findings reveal current open data indexes emphasize publication of data sets over the measurement of impact. The case study suggests that to be successful, cities should approach open data as a continuing journey and must actively engage other stakeholders, particularly intermediaries and citizens. Finally, we observe that common myths constructed around open data help promote open data at a strategic level, but must be viewed skeptically at the operational level.

1. Introduction

The movement toward open data, data “that is machine-readable, freely shared, used and built on without restrictions” [1], is gaining momentum. Open data is being driven by many perceived benefits, including increased government transparency, greater citizen participation, sustainable development, innovation, and economic development [2, 3]. However, many governments have yet to realize the promised benefits of open data [4, 5], leading some to suggest that certain myths surround open data [6, 7]. In truth, governments face many technical and organizational challenges to implementing and optimizing their open data programs [8, 9]. Moreover, open data success increasingly requires a collaborative effort between governments and their diverse stakeholders and intermediaries [10, 11].

A substantial amount of research has been devoted to understanding how open data initiatives evolve and

mature, such as the Open Data Maturity Model [12]. In addition, many indexes including the Global Open Data Index [13] and the Open Data Barometer [14] have been created to evaluate governments’ open data initiatives across a range of criteria. Although the criteria used by the various indexes are similar or overlapping, different approaches lead to different rankings [15]. The proliferation of measures creates ambiguity around the definition of a ‘top’ open data city and there is little research that assesses the appropriateness and value of the indexes themselves. Further, the mechanisms leading to the desired benefits of open data are still not well understood [16]. To the extent that open data indexes provide guidance to aspiring open data governments, examination of cities’ open data initiatives in relation to these indexes is important. Thus, our research questions ask: 1) what are the characteristics of a ‘leading’ open data city, and 2) how can cities improve their open data success?

To answer these questions, we conduct an exploratory case study of Edmonton, Alberta, Canada’s top ranked open data city. Using publicly available information, we investigate Edmonton’s top-ranking open data portal and then examine Edmonton’s evolution over the 8-year period from 2009 to 2017. By studying Edmonton’s open data experience, we uncover different factors and mechanisms contributing to its success. As such, our research offers contributions to both IS scholars and practitioners, from open data providers (e.g., cities) to users (e.g., businesses, citizens). Our findings extend the smart city and open data literature and provide tangible insights to cities undertaking open data initiatives.

In the next section, we present relevant theoretical background to the study. Then, we explain the research methodology. This is followed by the results, discussion, and conclusion.

2. Theoretical Background

The recent trend toward open data represents an extension of e-government and ‘smart city’ strategies [8] and is motivated in part by a vision of sustainable

development that balances economic, social, and environmental concerns. Open data promises a wide range of benefits for diverse stakeholders [6, 9, 16]. For governments, these benefits include reduced costs and operational efficiencies, economic development, greater citizen participation, improved public policies and services, and increased transparency. For citizens, open data can enhance employment and business opportunities, allow easier participation and collaboration with government, and better inform political decisions. Businesses benefit from open data through the development of new products, services, and business opportunities, while researchers should enjoy reduced costs of data acquisition, permitting greater innovation and knowledge creation.

Unfortunately, many of these anticipated benefits have yet to be realized [4, 5]. Anecdotal evidence also shows some basic failures, like the inability of governments to enforce open data repositories [17] and a lack of use of open data [18]. This disconnect between promise and reality suggests certain myths have been constructed around open data [6, 7].

Myths are dramatic narratives often beginning with real people or events that are subsequently embellished with fiction to provide explanations of some phenomenon [19]. Much debate exists around the value of myths. Proponents suggest myths are useful for providing explanations and cognitive structures to guide thinking, as well as establishing collective meaning and maintaining social solidarity and cohesion [19, 20]. In contrast, opponents argue the use of myths takes away from the building of theories and practices based on true experiences [20].

Six key myths of open data can be summarized as follows: 1) open data automatically and uniquely delivers benefits; 2) all public organization data should be published without restriction; 3) the main challenge of open data lies in the publication process; 4) open data can be used by everyone; 5) open data creates open and transparent governments; and 6) there is public interest in the reuse of open data [6, 7]. Collectively, these myths emphasize the potential benefits of open data, which may encourage the adoption of open data; however, they may also create unrealistic expectations for open data and mask major implementation challenges for cities [6, 7, 21].

2.1. Open Data Intermediaries

As noted above, a number of stakeholders expect to benefit from open data. However, the path to open data benefits is not a one-way street and, like other social-benefit innovations, requires collaboration [11] from a plurality of actors [22]. The fact that individual citizens (and cities) may not have the skills or knowledge to

fully exploit open data [6, 23], suggests intermediaries, such as application developers and researchers, are needed to take full advantage of open data [10, 11].

Intermediaries are a special type of stakeholder involved in supporting an innovation process by linking two or more actors in the innovation network [24]. Intermediaries usually play the role of knowledge providers [24]. This type of intermediary (also referred to as an infomediary) focuses on the collection and distribution of information and creates bridges between unconnected groups [24, 25]. In some cases, intermediaries play a more involved and interactive role, such as brokering a transaction between parties, setting standards, securing funding or support, or developing and implementing business strategies [24].

Gonzalez-Zapata and Heeks [26] propose that open data intermediaries perform five main roles: data demanders, data producers, data validators, application developers, and communicators. In other words, open data intermediaries can provide the knowledge and resources required to overcome barriers, such as absence of data quality and relevance, or lack of technical skills and resources to effectively use and manipulate data [26], leading to greater success.

2.2. Phases of Open Data Initiatives

Within both academia and practice, substantial effort has been devoted to defining and evaluating governments' progress with respect to open data [e.g., 27], with an underlying assumption that it takes time to achieve success. For example, research suggests that open data initiatives seem to rollout in waves [28]. First, governments focus on publishing existing data, modifying structures to facilitate open data, identifying key data sets, and making legislative changes to support these initiatives. Second, governments seek to improve the quality of open data and stimulate the use of open data. During the third wave governments look to enhance value by engaging users and external stakeholders to gain input and feedback and identify opportunities for improvement [28].

Building on this observation, more formal maturity models have been proposed. Maturity models outline a sequence of stages representing an anticipated, desired, or logical path from an initial state to maturity [29]. The 5-star open data deployment scheme was an early example (5stardata.info) related to technical dimensions of open data. More broadly, Lee and Kwak [30] proposed an Open Government Maturity Model with five stages. In this model, open data initiatives fall into the second maturity stage, data transparency, which is considered a prerequisite for more advanced stages involving open participation, collaboration, and engagement of citizens. Complementing this model,

Dodds and Newman [12] propose a five-themed Open Data Maturity Model in which a government's progress is measured on five main themes with each theme evaluated across five stages assessing activity level from initial ad hoc implementation of open data to the optimization of the open data program.

In parallel with the development of maturity models, numerous open data indexes have been created to rate governments' efforts. For example, for 2016 Canada was ranked second on the Open Data Barometer [14] and, Canada ranks fifth on the Global Open Data Index [13]. Within Canada, the Open Cities Index (OCI) was established in 2015 by Public Sector Digest (PSD) to benchmark cities. Like other indexes, the OCI draws from the extant literature and relies on self-reported data from municipalities which is then independently verified. For 2016, Edmonton, Alberta was Canada's top-ranked open data city [31].

Despite work in this area, important limitations remain that we hope to address with our research. First, due to the propagation of different measures there is a lack of clarity around the definition of open data success. Second, few studies empirically examine cities' performance in relation to existing indexes. A few studies [e.g., 15, 32] have looked at national or sub-national open data initiatives and indexes. However, contextual factors, including geographical level of the initiative, seem to influence open data success [33]. Finally, despite offering prescriptive frameworks, the extant research does not fully explain the underlying mechanisms by which cities can improve the success of their open data initiatives.

3. Methodology

This research involves an exploratory case study of one city's open data initiative. We purposefully chose Edmonton because it represents a unique case [34] as a leading open data city in Canada.

3.1. Research Site: City of Edmonton, Alberta

Edmonton is the capital of the province of Alberta and the fifth largest municipality in Canada. In 2016, Edmonton's metropolitan population was 1.3 million. The city is a major hub for the oil and gas industry and a major economic center for Alberta [35]. Edmonton officially launched its Open Data Catalogue in January 2010 with twelve datasets and continued to increase that number. More than 1,300 datasets were available in June 2017. As a result of its efforts, Edmonton claimed top spot in the 2015 and 2016 OCI rankings, giving it the title of the top open data city in Canada.

3.2. Data Collection and Analysis

To develop our case study, we reviewed the current status of Edmonton's open data initiative and its history from 2009 to June 2017. We examined three sources of publicly available data. First, we examined Edmonton's open data portal to analyze the data sets and tools available. We collected metadata about the data sets available (as of June 2017), including information such as file name, file type, number of views, tags of file, and last update date. Second, we collected information on the OCI and the criteria used to assess open data initiatives. Our objective was not to validate or repeat the evaluation process, but to understand the criteria used and Edmonton's top ranking. A full description of the OCI methodology can be found at [36]. Third, we collected relevant articles published in newspapers, journals, and magazines. Our decision was based on two factors: first, the media has been identified as a key infomediary that can shape how a new corporate practice is understood [37]; and second, we felt public data would better capture key historical events as compared to interviews because of personnel changes and memory distortions over time.

We identified potential articles using the Factiva's search engine and database, combining the keywords 'Edmonton' and 'open data', without any other restrictions. The search yielded 429 articles published between 2009 and 2017 in various outlets such as the Edmonton Journal, Postmedia, the Edmonton Sun, or Metro Canada. A total of 148 duplicates were automatically identified and removed by Factiva, leaving 281 papers for further analysis. Our next step was to ensure the relevancy of the articles by verifying that they provided information on Edmonton's open data initiative. One author performed the initial screening for inclusion. The procedure involved a rapid screening of the articles followed by a thorough examination to ensure the appropriateness of inclusion [38, 39]. In this step, we excluded 174 articles that were not published in English, were not related to both Edmonton and open data, or were duplicates not identified by Factiva. To validate the final set of articles, a second author reviewed the 107 remaining papers. A few disagreements were discussed between the two authors and consensus was reached on all articles. As a result, 12 additional papers were removed, leaving a final sample of 95 articles.

Figure 1 shows the distribution of the 95 articles according to publication date. There was a peak in 2010 when Edmonton launched its Open Data Catalogue. Between 2011 and 2014 there was a small increase in the number of articles published in newspapers, journals and magazines. The years 2015

and 2016 both saw a large increase in the number of articles. The number of articles in 2017 (6) reflects only articles published to May 2017 (5 months).

Our analysis of Edmonton's open data portal showed that all data available in June 2017 had an update date later than 2013, as shown on Figure 1. The majority of data sets were updated in 2015 and 2016 (742 and 540 datasets respectively).

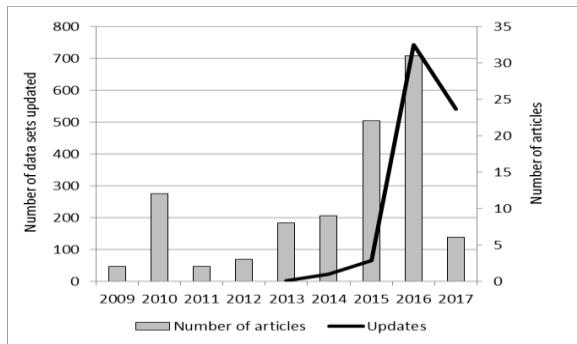


Figure 1. Articles and updates to data sets per year

Following identification of the articles, we proceeded with analysis, which involved descriptive coding [40] by two authors independently in parallel. We extracted and coded information related to the six myths of open data (see above), key events and actions that occurred during the implementation and evolution of Edmonton's open data initiative, stakeholders involved, and characteristics of data. We performed a thematic coding with emergent codes [40] of the articles using the data analysis software NVIVO. During the process, excerpts were regrouped under overarching themes related to the 'what' (i.e., content and format of data), 'when' (i.e., the events), 'how' (i.e., strategies and actions performed by the city and/or the community), and 'who' (i.e., the stakeholders) of Edmonton's open data initiative.

For each of the data sets evaluated in the OCI, we qualitatively assessed the relative visibility of different data sets (Table 1). For example, a data set was rated low if it received one or two passing references in the media, a rating of medium if there was some substantive reference to that particular data set, and a high rating if several articles in different years addressed different facets of the data set.

Next, our analysis involved the development of time-order matrices [40] in order to capture the change in elements over the 8-year period, spanning pre-launch of the Open Data Catalogue to Edmonton's current leadership position. To ensure reliability of the research, two authors independently coded the 95 articles and discussed their results until reaching a consensus view of Edmonton's open data initiative.

Finally, metadata on the open data sets were analyzed to complement our qualitative analysis. Due to space limitations, we present only confounding or inconsistent results from this analysis.

4. Results for RQ1: Characteristics of Edmonton as a Leading Open Data City

According to the OCI, a top open data city demonstrates its capabilities in three main areas: its *readiness* in terms of allocating sufficient financial and other resources (e.g., staffing), its ability to *implement* open data by publishing a variety of different data sets (32 data sets in 2016, see Table 1) in an open and accessible format, and its ability to generate positive economic and social *impacts* from the use of open data. Among these three capabilities, the OCI places substantially more weight, 68% of the total score, on the second capability. In contrast, a city's readiness and open data impact each account for only 16% of the total score. In effect, the OCI as currently formulated requires cities to show their capabilities for publishing data much more than demonstrating the final benefits of this open data publication.

Table 1. Data sets evaluated in OCI 2016

High visibility (7 data sets)	Property assessments, park inventory, public transit, real-time transit, service requests (311), crime statistics, web analytics
Medium visibility (5 data sets)	Council voting records, census data, traffic volumes, traffic accidents, road closures
Low visibility (14 data sets)	Government budget, election data, council expenses, municipal permits, zoning (GIS), base GIS data (roads, etc.), restaurant inspections, health performance, education performance, city services, bylaw infractions, environmental services, air quality, recreational programs
No mention (6 data sets)	Lobbyist information, public facilities and structures, company register, code enforcement violations, construction contracts, procurement contracts

In 2016, Edmonton earned a score of 98%, outpacing second-place city, Toronto, Ontario, at 76%. Edmonton scored 100% on both *readiness* and *impact* and 98% on *implementation* [41]. The OCI report explains that Edmonton's efforts to train the community (e.g., through hackathons) and adoption of official open data plans and policy contributed to its

readiness score [41]. In terms of *implementation*, “Edmonton has published online, up-to-date, freely available, machine-readable, automated datasets that are linked to APIs under almost all 32 categories of datasets” [41]. With respect to *impact*, Edmonton used a variety of tools to track metrics ranging from users’ use of published data to media coverage of open data initiatives [41]. Our own analysis of Edmonton’s open data portal shows varied use across data sets: the most popular had over 129,000 views, while the least viewed (6 in total) had no views. In total, Edmonton has received more than 1.5 million views of all open data sets with an average of 1140 views per data set.

For cities seeking to raise their open data performance, OCI clearly places the emphasis on making the data available before improving accessibility and impact. For instance, cities can earn points for simply making data available and keeping it updated (see Table 2). Additional points are awarded for data accessibility (e.g., standard formats and APIs). Various reasons may account for OCI’s weighting of implementation as compared to readiness and impact, not least of which is the inherent difficulty: “gauging the impact of a municipal open data initiative is one of the most challenging tasks for both municipalities and for those benchmarking initiatives” [36].

Table 2. OCI Points awarded per data set [36]

Each of the 32 datasets rated against 10 criteria (0 points for option 1; 1 point each for the other 10 options)	1.	No access to data
	2.	Data exists
	3.	Data is available in some form
	4.	Data is available in machine readable form
	5.	Data is accessible and permanent
	6.	Data is free
	7.	Data is available in bulk
	8.	Data is openly licensed
	9.	Data is up to date
	10.	Data is automated
	11.	Data is linked to API

5. Results for RQ2: Improving Open Data Success

In response to RQ2, our analysis revealed that Edmonton’s open data journey efforts evolved through four main phases in which two main dimensions – the data itself, both in content and format, and interaction with the broader community – contributed to success. Due to space limitations, a detailed chronology of all events cannot be presented herein so we limit our results to key elements that define each phase.

5.1. Phase 1: Creating Interest and the Open Data Vision

This phase occurred primarily in 2009 before the official launch of the Edmonton Open Data Catalogue. During this period, the city experimented with different ways of making data available to the public, such as in pdf form or through Google’s data standards. The city also consulted with diverse stakeholders to gather needs and expectations from the community. These efforts helped to create interest in open data and solidify Edmonton’s vision of using open data to improve transparency and overall management of the city: “*Much of the talk of open data’s potential played out on blogs and Twitter for months before data.edmonton.ca went live, giving citizens access to information including census data, bus stop locations and city council meeting schedules with the goal of improving government transparency and allowing us all to help the city run better*” [42]. This vision seems to have continued to motivate many of Edmonton’s actions over the next seven years.

5.2. Phase 2: Building the Catalogue and Engaging the Community

With the exception of special coverage around the launch of the open data catalogue, the 4-year period from 2010 to early 2014 was relatively quiet in the media. Through this phase, Edmonton continued to publish new data sets, attaining 415 by February 2014 and the city worked with other municipalities to adopt the Open311 format related to service requests.

A major contest was held in 2010 and annual hackathon events were organized to engage individuals and businesses in the community and encourage the development of value-adding open data applications. Despite the increasing number of data sets available, most public attention was focused on the use of open data in areas directly touching the lives of Edmontonians, such as transit, restaurant inspections, and property information. Although these apps may not have added much to government transparency, they validated the potential for open data. In addition, Edmonton recognized community engagement as a vital factor for success: “*we are getting ready to unveil the next generation Open Data Portal and the input we received from Edmonton’s vibrant Open Data community will help us build a user-friendly site that will allow users to capitalize on this opportunity*” [43]. At this time, journalists and application developers were the main users of open data in addition to serving as intermediaries for the general public.

5.3. Phase 3: The Community Takes Charge

The introduction of Edmonton's next generation Open Data Portal, including a Citizen Dashboard providing real-time data on the city's activities and responses to requests in late 2013 and the publication of new data sets in February 2014 marked a new phase in Edmonton's open data evolution. This phase was characterized by more direct and active involvement by the community that extended beyond the creation of applications. During this period, external stakeholders formed their own communities, initiated open data events in parallel with the city, and even collected and created new data sets.

Whereas phase 2 mostly involved efforts by the city to 'push' open data, phase 3 saw greater efforts from the community to 'pull' data from the city. Between 2014 and 2015, the number of open data sets grew to almost 700, covering a great variety of areas. The city added real-time transit information using GPS devices on buses, enhanced its mapping capabilities, and invested over \$200,000 in open data analysts. Meanwhile, the community group HackYEG organized its first hackathon in May 2014. The objectives and format of this event were largely the same as city-sponsored hackathons, however, the fact it was the first citizen-organized event represented a key milestone in Edmonton's open data evolution.

A second notable milestone was the city's willingness to accept data from non-government sources and citizens. The potential of crowdsourced data was demonstrated when a biking enthusiast, collaborating with other citizens, started collecting information on bike racks in the city in an effort to persuade the city to install more facilities. The creation of this data set, which spanned more than six months, caused the city to rethink how it would accept, publish and incentivize crowdsourced data because the quality of their open data was a substantial preoccupation. However, proponents of crowdsourced data argued that the value of such data outweighed the risks and that *"users will self-moderate the data"* [44] to correct problems, such as errors or intentionally false data.

5.4. Phase 4: Refining the Role of Intermediaries

The release of an updated version of the Open Data Catalogue in January 2016 marks the start of the fourth (and current) phase in Edmonton's open data evolution. With about 1000 data sets available at that time, most with machine-readable formats, APIs and data visualization options, the city clearly demonstrated both its commitment to and capability for publishing open data. Around this time, Edmonton

received the 2015 OCI top ranking of Canadian cities. From our analysis emerges an image of a city intensely proud of its open data achievements. Continuing its path of leadership, in early 2017, Edmonton became the first Canadian city to adopt the Open Data Charter.

In terms of engagement, the city was starting to reap benefits from its efforts to build a vibrant open data community, involving active and capable intermediaries (e.g., journalists, application developers, businesses). However, the release of the new Open Data Catalogue seems also to denote a shift in focus by establishing more direct links with citizens and open data end-users through disintermediation: *"the city has upgraded the user interface and functions to better suit online and mobile functionality. Users can now view open data in pre-categorized views, have access to apps that have been developed using the city's open data"* [45]. In conjunction with the new portal, Edmonton launched its Analytic Centre of Excellence's Open Analytics website providing *"tools to empower citizens to use open data to gain their own insights and features monthly updates such as step-by-step tutorials, project showcases and interactive data visualizations"* [45]. With these tools, individual citizens who do not necessarily have the skills necessary to build applications or process raw data can still make use of open data, thus reducing reliance on intermediaries. Further, city transparency is increased by putting these tools directly in the hands of citizens.

One somewhat unexpected finding from our analysis is that transparency does not seem to stop at the city: by combining open data with other data sources and building applications, greater visibility can be also gained in other sectors. For instance, two entrepreneurs in the real estate sector worked with a team for 18 months collecting and combining data to create an application to simplify the process of finding a home. They were motivated in part to bring increased transparency to an industry that was otherwise *"lacking transparency of knowledge and data"* [46].

Through 2016, the roles of the city and its stakeholders continued to evolve. One particular debate arose regarding the responsibility for building applications. External stakeholders (largely application developers) argued it would be more economical and beneficial for the city to focus on publishing data while allowing others develop creative applications. For its part, the city maintained that while it was *"not in the business of building apps"* [47] and had a preference for third-party solutions, it would build tools to respond to specific needs. For example, in 2016, the city owned and maintained seven applications, including tools for reporting crimes, reserving library books, and paying for parking.

A second aspect of this role redefinition was the increasing use of open data by the city itself. Having worked to make an enormous amount of data available to external stakeholders, Edmonton began to recognize (starting in phase 3 and growing in phase 4) the value of this data and make use of it in the operational and strategic management of the city, from determining health care needs to developing better approaches to policing and safety.

6. Discussion

We chose to study Edmonton as an exemplar case to better understand the mechanisms and processes contributing to open data success. With a single case study, firm conclusions are impossible, however, three main insights emerge from our work: first, open data should be viewed as a continuing journey; second, cities cannot achieve open data success on their own and must involve other stakeholders; third, open data myths may help promote open data at a strategic level, but must be viewed skeptically at the operational level.

6.1. The Open Data Journey

When it comes to open data, what constitutes success and what is the path to success? These questions motivated our research and are of high importance to practice. One way to measure success is the use of indexes. To the extent that such third-party evaluations evaluate a city's open data performance, they also provide an indication of a city's maturity. The 2016 OCI report, for example confirms "that Canada's municipalities are on the right track to open data maturity" [41]. In constructing its measure, OCI has taken an approach consistent with the waves of open data implementation [28]. In its scoring, the greatest weight is given to publishing specified data sets and making the data accessible in diverse formats. While Edmonton received near perfect scores, other Canadian cities did not, suggesting that open data remains a challenge for many cities.

In this light, we suggest the Edmonton experience provides an alternative view of open data success. Open data is not a process with a definitive point of maturity to be achieved through a sequence of stages in a logical path. Instead, open data is a continuing journey as cities navigate through changing social, technical and data landscapes. Edmonton began with a clearly articulated vision for open data supported at all levels from the mayor to IS department. Open data was a key part of the city's strategy, not simply another IS project done in response to a current fad. Our analysis also shows that once the vision was firmly established, Edmonton then focused not only on publishing data,

but also building the open data community. The city seemed to understand that making data available would not be sufficient for realizing substantial long-term value. As a result, the city took efforts to ensure it was "*open by default*" [48], participated in standards definition, consulted with its community, and adopted new technologies, such as mapping, visualization, and analytics, to support the use of open data.

By the standards of OCI, Edmonton would be considered a more mature open data city. However, the city could do more to augment its open data initiative. With over 1300 open data available, the city appears to have mastered the challenges of publishing its data in open and accessible formats. The data has become, arguably, less critical to open data success, and has been replaced by the dynamic and changing relationships between the city and its stakeholders as the critical element of success.

6.2. City, End-users, and Intermediaries

The second insight we draw from Edmonton's experience is the involvement of various stakeholders as intermediaries and their evolving roles throughout the open data initiative. Here, our findings are consistent with the five roles suggested for intermediaries in open data initiatives [26]: demanders, producers, validators, developers, and communicators of data. Not only was each role adopted by one or more intermediary, but the intermediaries also evolved in their functions and performed different tasks over the phases of Edmonton's open data journey.

In the role of data **demanders**, intermediaries convey public opinion and requirements as they request particular data sets or promote the adoption of policies [26]. This was the first role assumed by Edmonton's intermediaries and the city's openness community input continues today. In 2009, even before the launch of its first open data sets, the city sought input from the community, consulting potential intermediaries and users to gather their needs and expectations. Once the Open Data Catalogue was launched and new data sets published, additional intermediaries emerged and motivated the release of additional data from the city. For example, citizens and the media requested specific datasets and activists put pressure on the city to adopt standards, policies, and agreements with other public agencies. Operating in this role, the intermediaries helped to ensure, from the beginning, the relevance and quality of published data.

As **producers**, intermediaries collect new data or combine existing data in order create new open data sets [26]. In 2015 (phase 3), a notable shift in the provision of data occurred when Edmonton started to accept data from external sources, including citizens or

businesses. The integration of crowdsourced data enabled both end-users and intermediaries to play the role of producers of open data in collaboration with the city. This trend continues as Edmonton has started working with other agencies, such as Alberta Health Services, to co-create and publish new data sets.

Availability of the data is not sufficient to realize benefits of open data. **Validators** contribute to open data success by checking available data and confirming its usability and relevance [26]. With the exception of evaluating whether the data sets are up-to-date, the OCI index does not explicitly measure the quality of open data (see Table 1). However, our evidence suggests data quality and relevancy was a concern for Edmonton. We also note that journalists, developers, academics, and citizens at various times played the role of validators. For example, inputs from the community were used to guide and validate the development and provision of services, from the Open Data Portal to the Citizen Dashboard and the Open Analytics website. Additionally, with the arrival of crowdsourced data users were expected to monitor the accuracy and usability of the data. The contributions of intermediaries in this role helped improve the quality and relevance of open data and related tools.

Developers contribute to the accessibility and usability of open data by creating websites and applications processing open data [26]. Like the role of demanders, the role of developer was one of the first to be assumed by intermediaries. Shortly after the release of Edmonton's first data sets, tech-savvy individuals and companies started to develop different applications for open data. These applications demonstrated the possibilities for open data and helped engage the community by transforming open data into meaningful information. Developers enabled ordinary citizens and the city itself to overcome barriers related to technical skills and resources. Over time, the developer role evolved into a joint effort, as both the city and intermediaries organized hackathon events and developed applications. This role was highly instrumental in the success of Edmonton's open data initiative because developers provided citizens ways of using and making sense of raw data.

In the role of **communicators**, intermediaries popularize open data by translating and communicating complex raw data sets [26]. In Edmonton, multiple stakeholders, such as the media, academics, and businesses, fulfilled the role of communicators by combining data sets from different sources, translating data into meaningful information, interpreting complex information, and diffusing information and knowledge to end-users. In one case, for example, a real estate company created interactive maps showing housing-related patterns. Besides individual applications and

data sets, hackathons facilitated the job of communicators, allowing participants to explore raw data, make sense of the data, and find use for the data.

In this section, we described the roles played by stakeholders as if they were distinct from each other. In reality, roles were not defined explicitly at the beginning and constantly evolved. For example, end-user citizens became more proactive over time and open data activists assumed the role of intermediaries. Tech-savvy individuals and businesses also wore many hats, increasing interaction, collaboration, and sometimes challenging the city's decisions. For its part, the city seemed to take a pragmatic approach, adapting to emerging situations in order to achieve its vision for open data and transparency.

6.3. Myths of Open Data

Our third insight concerns the myths surrounding open data. As discussed earlier in the paper, myths may both help and hinder the adoption of different IS, an observation that seems to bear out in the case of Edmonton. Initially, our analysis seemed to suggest the six myths of open data were of little consequence to Edmonton's open data initiative. However, when we delved deeper into the data, we observed that the myths can apply at two levels. Most of the myths (1, 2, 3, 4, and 6) relate to the operational details of open data: the challenges of publishing and using data, while the fifth myth (related to transparency), seems to apply, at least in the case of Edmonton, at the strategic level.

Strategically, Edmonton drew heavily on the idea that open data creates transparent governments to establish the vision for open data and to justify actions: *"open data is a way of telling the good, the bad and the ugly of what's happening in Edmonton [48]"*. In contrast, our analysis suggests Edmonton did not get caught by the other myths, particularly those related to the challenges of publication, automatic value, public interest, and user capabilities. Instead, Edmonton appears to have tackled these myths head-on, using mechanisms (such as building the open data community) to ensure success. Whether consciously or not, the city seems to have adopted a pragmatic view of the myths of open data: it used the myths strategically to advance the initiative while at the same time assuring the myths did not impede progress.

7. Conclusion

Our research offers three main contributions to research and practice. First, we provide rich description of an exemplar case, Edmonton, and in so doing illustrate the intricacies of a successful open data initiative. Our research sheds new light on the myths

and realities of open data and extends theory regarding the role of intermediaries by revealing the evolutionary processes and dynamic relationships between cities, end-users, and infomediaries. Second, our results complement previous maturity models by highlighting the importance of adopting long-term approaches for open data initiatives. In Edmonton, for example, it took about four years for the city to begin to realize substantial benefits and start to refine the roles of stakeholders within the open data community. This need for long-term commitment cannot be underemphasized, particularly as cities operate in unstable political environments with inherently short-term goals. Third, our results provide direction to research and practice by revealing the need for more sophisticated measures of open data maturity and performance. Existing indexes provide a good starting point, but these are still relatively blunt instruments that may not be capable of taking into account non-linear paths to open data success.

There are certain limitations of this work, primarily arising from our single case study approach and use of public data. With a single case study, broad generalizations are not possible as the findings are specific to the study context. To address this concern, future research involving other cities similar to Edmonton would allow for comparison and the development of grounded theory. With respect to the data sources, we did not interview people involved directly in the Edmonton open data initiative. Collecting 'insider' data would permit data triangulation leading to more robust research findings. Primary data would also provide a view into the motivations, policies, and behind-the-scenes mechanisms leading to certain events. Interviews with key participants could shed light on the challenges faced in implementing open data that cannot be answered fully using secondary data. Finally, our research looked only at the OCI, although other indexes exist. Future research could examine the validity and reliability of the OCI and other measures more thoroughly. As an adjacent stream of research, we suggest researchers work in collaboration with practitioners to develop more sophisticated approaches for measuring the impact of open data, such as those used for assessing impact in online communities [49].

Open data has the potential to deliver many benefits to cities and all their stakeholders. We hope the new knowledge created through this research will provide a stepping-stone to realizing that potential.

8. Acknowledgement

We thank Mathieu Marcotte for his contribution to this research.

9. References

- [1] Government of Canada, "Open Data 101", 2017, Available: <http://open.canada.ca/en/open-data-principles>
- [2] Hossain, M.A., Dwivedi, Y.K., and Nripendra, P.R., "State-of-the-Art in Open Data Research: Insights from Existing Literature and a Research Agenda", *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 2016, pp. 14-40.
- [3] Smith, F., Gerry, W., and Truswell, E., *Supporting Sustainable Development with Open Data*, Open Data Institute, 2015.
- [4] Agrawal, D., Kettinger, W.J., and Zhang, C., "The Openness Challenge: Why Some Cities Take It on and Others Don't", *Twentieth Americas Conference on Information Systems*, 2014, pp. 1-7.
- [5] Ham, J., Lee, J.-N., Kim, D.J., and Choi, B., "Open Innovation Maturity Model for the Government: An Open System Perspective", *Thirty Sixth International Conference on Information Systems*, 2015
- [6] Janssen, M., Charalabidis, Y., and Zuiderwijk, A., "Benefits, Adoption Barriers and Myths of Open Data and Open Government", *Information Systems Management*, 29(4), 2012, pp. 258-268.
- [7] Hellberg, A.-S., and Hedstrom, K., "The Story of the Sixth Myth of Open Data and Open Government", *Transforming Government: People, Process and Policy*, 9(1), 2015, pp. 35-51.
- [8] Bakici, T., Almirall, E., and Wareham, J., "A Smart City Initiative: The Case of Barcelona", *Journal of the Knowledge Economy*, 4(2), 2013, pp. 135-148.
- [9] Cranefield, J., Robertson, O., and Olivier, G., "Value in the Mash: Exploring the Benefits, Barriers and Enablers of Open Data Apps", *European Conference on Information Systems*, 2014
- [10] Young, M., and Yan, A., "Civic Hackers' User Experiences and Expectations of Seattle's Open Municipal Data Program", *50th Hawaii International Conference on System Sciences*, 2017, pp. 2681-2690.
- [11] Chan, C.M.L., "From Open Data to Open Innovation Strategies: Creating E-Services Using Open Government Data", *46th Hawaii International Conference on System Sciences*, 2013, pp. 1890-1899.
- [12] Dodds, L., and Newman, A., "Open Data Maturity Model", 2015, Available: <http://theodi.org/guides/maturity-model>
- [13] Open Knowledge International, "Place Overview", 2017, Available: <http://index.okfn.org/place/>
- [14] World Wide Web Foundation, "Open Data Barometer", 2017, Available: http://opendatabarometer.org/?_year=2016&indicator=ODB
- [15] Lnenicka, M., "An in-Depth Analysis of Open Data Portals as an Emerging Public E-Service", *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 9(2), 2015, pp. 589-599.
- [16] Parycek, P., Höchtl, J., and Ginner, M., "Open Government Data Implementation Evaluation", *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 2014, pp. 80-99.

- [17] Eaves, D., "Canada's Opaque Transparency – an Open Data Failure", 2014, Available: <http://eaves.ca/2014/03/04/canadas-opaque-transparency-an-open-data-failure/>
- [18] Worthy, B., "Open Data on Council Spending Is Largely Unread by Voters", 2015, Available: <http://www.theguardian.com/public-leaders-network/2015/jun/10/council-finance-open-data-unread-by-voters>
- [19] Hirschheim, R., and Newman, M., "Symbolism and Information Systems Development: Myth, Metaphor and Magic", *Information Systems Research*, 2(1), 1991, pp. 29-62.
- [20] Abel, C.F., "Why Myth Could Matter (but Probably Shouldn't) in Public Administration", *International Journal of Organization Theory and Behavior*, 14(2), 2011, pp. 168-199.
- [21] Bekkers, V., and Homburg, V., "The Myths of E-Government: Looking Beyond the Assumptions of a New and Better Government", *Information Society*, 23(5), 2007, pp. 373-382.
- [22] Corbett, J., and Montgomery, W., "Environmental Entrepreneurship and Interorganizational Arrangements: A Model of Social-Benefit Market Creation", *Strategic Entrepreneurship Journal*, In press.
- [23] Martin, C., "Barriers to the Open Government Data Agenda: Taking a Multi-Level Perspective", *Policy & Internet*, 6(3), 2014, pp. 217-240.
- [24] Howells, J., "Intermediation and the Role of Intermediaries in Innovation", *Research Policy*, 35(5), 2006, pp. 715-728.
- [25] Deephouse, D.L., and Heugens, P.P.M.A.R., "Linking Social Issues to Organizational Impact: The Role of Infomediaries and the Infomediary Process", *Journal of Business Ethics*, 86, 2009, pp. 541-553.
- [26] Gonzalez-Zapata, F., and Heeks, R., "Understanding Multiple Roles of Intermediaries in Open Government Data", 13th International Conference on Social Implications of Computers in Developing Countries, 2015, pp. 1-10.
- [27] Ubaldi, B., "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives": OECD Working Papers on Public Governance, No 22, OECD Publishing, Paris, 2013
- [28] Nugroho, R.P., Zuiderwijk, A., Janssen, M., and De Jong, M., "A Comparison of National Open Policies: Lessons Learned", *Transforming Government: People, Process and Policy*, 9(3), 2014, pp. 287-308.
- [29] Poeppelbuss, J., Niehaves, B., Simons, A., and Becker, J., "Maturity Models in Information Systems Research: Literature Search and Analysis", *Communications of the Association for Information Systems*, 29(Article 27), 2011, pp. 505-532.
- [30] Lee, G., and Kwak, Y.H., "An Open Government Maturity Model for Social-Media Based Public Engagement", *Government Information Quarterly*, 29, 2012, pp. 492-503.
- [31] Public Sector Digest, "Open Cities Index", 2017, Available: <http://publicsectordigest.com/open-cities-index-oci>
- [32] Bello, O., Akinwande, V., Oluwatoyosi Jolayemi, O., and Ibrahim, A., "Open Data Portals in Africa: An Analysis of Open Government Data Initiatives", *African Journal of Library, Archives and Information Science*, 26(2), 2016, pp. 97-106.
- [33] Susha, I., Zuiderwijk, A., Charalabidis, Y., Parycek, P., and Janssen, M., "Critical Factors for Open Data Publication and Use: A Comparison of City-Level, Regional, and Transnational Cases", *eJournal of eDemocracy and Open Government*, 7(2), 2015, pp. 94-115.
- [34] Dubé, L., and Paré, G., "Rigor in Information Systems Positivist Case Research: Current Practices, Trends and Recommendations", *MIS Quarterly*, 27(4), 2003, pp. 597-635.
- [35] Wikipedia, "Edmonton", 2017, Available: <http://en.wikipedia.org/wiki/Edmonton>
- [36] Public Sector Digest, "Open Cities Index Methodology 2016", 2017, Available: <http://publicsectordigest.com/open-cities-index-methodology-2016>
- [37] Grafstrom, M., and Windell, K., "The Role of Infomediaries: CSR in the Business Press During 2009-2009", *Journal of Business Ethics*, 103, 2011, pp. 221-237.
- [38] Kitchenham, B., and Charters, S., "Guidelines for Performing Systematic Literature Reviews in Software Engineering", 2007, Available: <http://community.dur.ac.uk/ebse/resources/guidelines/Systematic-reviews-5-8.pdf>
- [39] Templier, M., and Paré, G., "A Framework for Guiding and Evaluating Literature Reviews", *Communications of the Association for Information Systems*, 37(Article 6), 2015, pp. 112-137.
- [40] Miles, M., B., Huberman, A.M., and Saldana, J., *Qualitative Data Analysis: A Methods Sourcebook*, SAGE Publications, Inc., Thousand Oaks, CA, 2014.
- [41] Public Sector Digest, "2016 Open Cities Index - Top 20 Results", 2017, Available: <http://publicsectordigest.com/2016-open-cities-index-top-20-results>
- [42] Fong, J., "Reason 29: Engaged Online Communities", *Edmonton Journal*, January 23, 2010.
- [43] Anonymous, "Canada: Minister Clement Taps Edmonton Expertise at Open Data Roundtable", *Mena Report*, April 4, 2013.
- [44] Johnson, D., "YEG Treemap Tracking Town's Timber", *Edmonton Examiner*, May 7, 2014.
- [45] Ostad, R., "City Updates Open Data Platform", *Edmonton Examiner*, January 20, 2016.
- [46] Mckague, J., "Data Aid for Homebuyers; User-Friendly Tool Filters Online Searches from Crime Rates to Commute Distance", *Edmonton Journal*, March 5, 2016.
- [47] Clancy, C., "Software Developer Says City of Edmonton Apps Miss the Mark", *Postmedia Breaking News*, September 15, 2016.
- [48] Stolte, E., "City Open-Data Policy Leads to Efficiencies", *Edmonton Journal*, December 7, 2016.
- [49] Cuellar, M., Young, B., Sainsbury, R., and Takeda, H., "Online Community Influence: A Study Using the Hirsch Metric and Social Network Analysis", *America's Conference on Information Systems*, 2013